Report No. UT-18.xxx

THE IMPACT OF DEMOGRAPHICS ON CRASH RISK IN DIVERSE UTAH COMMUNITIES

Prepared For:

Utah Department of Transportation
Research Division

Submitted By:

ACTIVE PLANNING

Authored By:

Shaunna K. Burbidge, PhD

Draft Final Report
December 2019

RESEARCH
UDOT Keeping Utah Moving
Utah Department of Transportation - Research Division
4501 South 2700 West - P.O. Box 148410 - SLC,UT 84114-8410

DISCLAIMER

The authors alone are responsible for the preparation and accuracy of the information, data, analysis, discussions, recommendations, and conclusions presented herein. The contents do not necessarily reflect the views, opinions, endorsements, or policies of the Utah Department of Transportation or the U.S. Department of Transportation. The Utah Department of Transportation makes no representation or warranty of any kind, and assumes no liability therefore.

ACKNOWLEDGMENTS

The authors acknowledge the Utah Department of Transportation (UDOT) for funding this research, and the following individuals from UDOT on the Technical Advisory Committee for helping to guide the research:

- Vincent Liu- UDOT Research
- Travis Evans- UDOT Traffic and Safety
- Jeff Lewis- UDOT Traffic and Safety
- Carrie Silcox- UDPS Highway Safety Office
- Heidi Goedhart, UDOT Planning

Additionally, the project team would like to acknowledge Evan Curtis from the Utah Governor's Office of Management and Budget, Pam Perlich, Director of Demographic Research at the Kem C. Gardner Policy Institute; and Scott Festin, of the Wasatch Front Regional Council for their assistance in providing data for this project.

TECHNICAL REPORT ABSTRACT

1. Report No. UT- 18.XXX	2. Government Accession No. N/A	3. Recipient's Catalog No. N/A
4. Title and Subtitle The Impact of Demographics on Crash Risk in Diverse Utah Communities		5. Report Date October 2019 6. Performing Organization Code
7. Author(s) Shaunna K. Burbidge, PhD; David Bassett, PE		8. Performing Organization Report No.
9. Performing Organization Name and Address Active Planning PO Box 393 Kaysville, UT 84037		10. Work Unit No. 5H0…, 8RD… (PM will provide) 11. Contract or Grant No. 13… (PM will provide)
12. Sponsoring Agency Name and Address Utah Department of Transportation 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410		13. Type of Report & Period Covered Draft Final Feb 2013 to Feb 2015 14. Sponsoring Agency Code PIC No. (PM will provide)
15. Supplementary Notes Prepared in cooperation with the Utah Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration		

16. Abstract

Demographics have long been shown to impact crash risk, particularly gender, age, race/ethnicity, employment, educational attainment, household size and even marital status. However, to date, there has been no thorough analysis of how these dynamics may be at play in Utah. Rather than looking at areas with populations that may exhibit higher risk factors for certain types of crashes, safety analysis has focused more on which locations experience larger numbers of crashes. This research evaluates how different crash types cluster across demographic groups and different geographic areas based on population characteristics and identifies risk factors and criteria for specific populations or areas where populations are concentrated. This study utilized traffic crash data from UDOT's crash database (Numetric/Safemap), and the Fatality Analysis Reporting System (FARS) for 215,000 crashes occurring in Salt Lake County from 2010-2018. Additionally, demographic data was collected from the U.S. Census and the American Community Survey for 212 census tracts in Salt Lake County. Using a combination of Binary Logistic Regression Models, crash types and manner of collision were correlated to local demographics. Several demographic characteristics were significantly correlated to the likelihood of severe crashes occurring in a given census tract. It is recommended that UDOT undertake a thorough evaluation of demographic trends in areas with higher frequencies of crashes in addition to traditional geometric design and built environment analyses. This will ensure that all appropriate measures are considered, and that suitable educational and informational campaigns can be implemented in addition to engineered countermeasures.

17. Key Words Demographics, crash risk, ethnicity, risky behavior, communication, population groups, factor analysis, population patterns		18. Distribution Statement Not restricted. Available through: UDOT Research Division 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410 www.udot.utah.gov/go/research		23. Registrant's Seal N/A
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 70	22. Price N/A	

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

ACS	American Community Survey
CATI	Computer Assisted Telephone Interview
CAPI	Computer Assisted Personal Interview
CBD	Central Business District
DPS	Department of Public Safety
DUI	Driving Under the Influence
FHWA	Federal Highway Administration
LEO	Law Enforcement Officer
NHTSA	National Highway Traffic Safety Administration
PSL	Posted Speed Limit
SE	Socio-Economic
TAZ	Traffic Analysis Zone
UDOT	Utah Department of Transportation
WHO	World Health Organization

EXECUTIVE SUMMARY

Demographics have long been associated with crash risk. Gender, age, race/ethnicity, employment, educational attainment, household size and even marital status have all been correlated with different types of crashes. However, to date, there has been no thorough analysis of how these dynamics may be at play in Utah. Rather than looking at areas with populations that may exhibit higher risk factors for certain types of crashes, safety analysis has focused more on which locations experience larger numbers of crashes. This research seeks to evaluate how different crash types cluster across demographic groups and different geographic areas based on population characteristics, in order to provide context sensitive safety solutions. Additionally, risk factors are identified for specific populations or areas where populations are concentrated, and specific populations are identified which warrant special attention during planning or implementation.

Data collection was focused on Salt Lake County, Utah and included 215,645 total crashes, which occurred from 2010-2018. All crash data was compiled using the Utah Safemap tool by Numetric. Demographic data was compiled for the 215 census tracts located within the county using a combination of U.S. Census, American Community Survey, and WFRC Socio-economic data. All data was collected electronically.

Several Binary Logistic regression models were employed to examine relationships between crash types and demographic characteristics. For these analyses the most dangerous crash types were identified including: driving under the influence (DUI), crashes involving older drivers, collisions with a fixed object, unrestrained crashes, disregard for traffic control devices, and single vehicle crashes. In order to ensure that the analysis provided was robust and comprehensive, many additional variables were included in various iterations of reach model to ensure that latent confounding variables did not skew the analysis and that collinearity was reduced or eliminated.

Numerous demographic characteristics were correlated to the likelihood of specific crash types occurring within a given census tract. DUI crashes were most likely to occur in areas with a higher percentage of residents with a bachelor's degree, higher rates of biking to work. These crashes were also more likely to occur in areas with a low percentage of teens, Hispanics, and smaller households.

Rashes involving older drivers were more likely to occur in areas with high percentages of young adults (ages 19-24) and Seniors. These crashes were also more likely to occur in areas where the population has a bachelor's degree or less. Areas with a high percentage of Asians, Native Americans, Polynesians and Hispanics were significantly less likely to have crashes involving older drivers.

Crashes resulting from a collision with a fixed object were most likely to occur in areas with a highly educated population (bachelor's, graduate or professional degree). These crashes were less likely to occur in areas with a large teen population. Additionally, tracts with a large amount of walking were likely to see fewer of these crashes while areas with a large percentage of auto commuters were significantly more likely to experience these crashes nearby. Unrestrained crashes were most likely to occur in areas with a high percentage of walkers and bikers. Areas with a large population of Native Americans were also significantly more likely to experience unrestrained crashes nearby.

Crashes involving the disregard of a traffic control device were most likely to occur in areas with a larger population of older adults (ages 35+) and areas with lower educational attainment. Areas with higher unemployment were significantly less likely to experience these types of crashes. Single vehicle crashes were significantly less likely to occur in areas with higher transit ridership, and walking to commute, and more likely to occur in highly educated areas.

It is recommended that UDOT undertake a thorough evaluation of demographic trends in areas with higher frequencies of crashes in addition to traditional geometric design and built environment analyses. This will ensure that all appropriate measures are considered, and that suitable educational and informational campaigns can be implemented in addition to engineered countermeasures.

1.0 INTRODUCTION

1.1 Problem Statement

While Utah is often seen as having a rather homogenous population, demographic diversity does exist, and for some variables is even clustered in specific areas and communities. The literature is rife with examples of how demographics impact crash risk, particularly as they relate to gender, age, race/ethnicity, employment, educational attainment, household size and even marital status. However, to date, there has been no thorough analysis of how these dynamics may be at play in Utah. Rather than looking at areas with populations that may exhibit higher risk factors for certain types of crashes, safety analysis has focused more on which locations experience larger numbers of crashes. Most of the local research has examined transportation network characteristics and geometric design or built environment features and how they may be creating conditions which result in more crashes. Demographics and population characteristics, for the most part, have been overlooked as predictors of crash risk or rates, and have been used only as descriptive variables or included as statistical controls. The missing piece in understanding traffic safety and crash risk is understanding how different types of crashes are spread across different groups of people and different communities in the state.

According to the World Health Organization (WHO) traffic crashes accounted for 1.35 million fatalities in 2016, the 8th leading cause of death in all age groups around the world (World Health Organization, 2018). By 2020 it is predicted that the number of traffic fatalities will surpass 2 million as countries become motorized faster. The most predominant factors used for predicting crash risks and rates can be split into three categories: environmental, vehicle, and human, the latter of which is often over looked even as we try to reduce both crashes and fatalities.

1.2 Objectives

The purpose of this research is to examine relationships between demographics and crash risk and rates in Salt Lake County, Utah. By examining gender, age, ethnicity, income and education of census tracts in Salt Lake County and correlating these characteristics against crash types and crash severity within specific cluster we will be able to determine which demographic variables prove reliable for identifying locations that are most at risk. This research has two main goals. First, we will evaluate how different crash types cluster across demographic groups and different geographic areas based on population characteristics, in order to provide context sensitive safety solutions. Second, we will identify risk factors and criteria for specific populations or areas where populations are concentrated, as well as identifying specific populations which warrant special attention during planning or implementation.

1.3 Scope

This study utilized traffic crash data from UDOT's crash database (Numetric/Safemap), and the Fatality Analysis Reporting System (FARS) for all crashes occurring in Salt Lake County from 2010-2018. Additionally, demographic data was collected from the U.S. Census and the American Community Survey for 212 census tracts in Salt Lake County. Crash types and criteria were correlated to local demographics using as series of quantitative data models. Finally, the research determined whether crashes involved local traffic or through traffic based on the home address of the individuals involved in each crash.

1.4 Outline of Report

The report is organized into six sections, as follows: Section 2 provides a brief literature review examining existing research on specific demographic characteristics and their impact on crash risk and different types of crashes. Section 2 also includes a description of the study methods and justifications. Section 3 presents the study data collected and provides summary characteristics for the crash reports. Section 4 presents both qualitative and quantitative analysis of the observed non-motorized travel behavior. Section 5 provides conclusions based upon the data analysis, and Chapter 6 outlines the author's recommendations for implementation.

2.0 RESEARCH METHODS

2.1 Overview

A thorough literature review examined demographics and their relationship to traffic crashes. This chapter provides background information on the impacts of gender, age, race and ethnicity, income, education, and driver's license possession. It also includes a discussion of the research methods employed and the justification for each.

2.2 Literature Review

The literature on motor vehicle crash risk predominantly focuses on either existing problems with the roads and how to make them safer, or the vehicles involved. Because of the focus on environment and vehicles, a comparatively small amount of research has been conducted on the behavioral component. The literature included in this review go into depth on specific demographics rather than an overview of many. Therefore, each piece of literature covered in this review focuses on no more than two of the demographics outlined by our research.

2.2.1 Gender

The literature shows that males are disproportionately involved in more vehicle accidents, with WHO reporting around three quarters (73%) of fatalities being male. Research shows that young males are most at risk in traffic crashes due to driver error. These violations then decrease with age (De Winter and Dodou, 2010). While females commit more unintentional errors in their driving (Oppenhiem et al, 2016) it seems males are at a higher risk of risk-taking behavior. Other aspects that play into the gender disproportionality are psychological factors that arise with gender roles and the more masculine and seemingly apathetic perception males often have. Males between the ages of 21-30 exhibit the highest number of driving faults stemming from negligence (Karacasu and Er, 2011). And while this does decreases with age, a similar pattern is not observed in females. However, other literature suggests that once annual mileage is considered, the impact of gender almost disappears (Lourens, Vissers, Jessurun, 1999) suggesting that female crashes occur less, due to their lower annual mileage. The same research found that age is the only variable, considering age, sex and education, that stays significant when correcting for annual mileage.

2.2.2 Age

Traffic crashes are the leading cause of death for individuals between the age of 5-29 (WHO, 2018). A lot of the research on age is closely tied to gender, as shown previously with young males being most at risk. As a driver gains more experience on a road, their performance improves meaning errors made on exams such as the Driver Behavior Questionnaire (DBQ) reduce with age. When looking at stop sign violations and those at fault in a traffic crashes due to a failed stop, the two largest groups involved are those under 18 or over 64 (Retting, et al, 2003). Research also points to a younger population as contributing greatly to drunk driving crashes. While individuals between the ages of 16 and 25 make up 16% of the U.S. population the same population accounts for 28% of all fatalities in drunk driving incidents (Larimer, et al, 1998). Data shows that lack of experience paired with the misuse of alcohol is a large problem with youth in this demographic.

2.2.3 Race and Ethnicity

Race and ethnicity patterns can be seen when examining the demographics of those who are most at risk of being involved in crashed. The National Highway Traffic Safety Administration (NHTSA) has published data from 1999-2004 identifying race and ethnicity in fatal vehicle crashes. Fatalities are highest in Native American drivers as a percentage of all deaths within their ethnic backgrounds, independent from socioeconomic factors (Roudsani, Ramisetty-Mikler, and Rodriguez, 2009). Fatally injured drivers who had been drinking was also highest for Native Americans (57%) and were also less likely to hold a valid US driver’s license or wear a seatbelt. These statistics are nationwide which may differentiate this study from other results. Many these fatalities occur on rural roads, particularly on Native American Reservations where roads may not be maintained as well as those in urban developments, and DUI enforcement is more limited. Rural roads may also contribute more fatal crashes due to the time it takes emergency responders to arrive on the scene. There are some similarities when this nationwide data is compared to smaller scale research examining race and ethnicity in Arizona (Campos-Outcalt, Bay, Dellapena and Cota, 2003). Again, Native Americans were the only ethnic group to have consistently higher fatalities in all age and sex categories. Despite this, when Native American's motor vehicle fatalities were corrected for their urban rate their numbers fell within the norm.

NHTSA data also suggests that the Hispanic population may be over represented in their data as many results are like those observed in the Native American community data. Hispanic populations have the second highest rate of fatalities, fatal crashes while driving under the influence, and seatbelt non-use. Research conducted on stop sign violations (Romano, Voas, Tippetts, 2006) suggested that a large reason the Hispanic population could see more traffic deaths may be due to confusion in signal and signage understanding, and a struggle with the rules and regulation in the United States. Failing to follow traffic laws is prevalent in Latin America. A study in Argentina found that in a single day in Buenos Aires, over 794,000 failures to comply with stop signs occurred (Beltramino and Carrera, 2007). However, the same Arizona study showed a different picture of Hispanic drivers. In comparison with Non-Hispanic White drivers, Hispanic drivers were found to have significantly lower fatality rates in all categories with the exception of urban males, suggesting that gender is more of a predictor than ethnicity.

2.2.4 Income

Income inequality has been related to crash fatalities, and in international studies the lowest income countries have not only the highest mortality rates but also the highest morbidity burden on their economic development (Laflamme and El-Khatib, 2018). While some of the factors that would have affected this result are not relevant to this study, as infrastructure and policies can differ greatly from country to country, aspects such as road behavior and means of transport can be relevant. Income inequality in an area can create a heterogeneity of vehicles with the wealthy buying larger and typically safer cars while lower income households often rely on what they can afford, resulting in higher rates of non-motorized modes such as walking and cycling, which can be more dangerous. While this in and of itself does not increase the frequency of traffic collisions, it does affect the fatality rates and has impacted lower income populations the most. For every 1% increase in vehicle weight, fatality risk is reduced by 5% (Anbarci, Escaleras and Register, 2009) meaning as the wealthy increase the size and weight of their cars, further protecting the occupants of their vehicles, it concomitantly increases the injury risk of those using other transportation modes. Other research has found strong correlations between unemployment rates and traffic mortality, once again impacting those who are less advantaged. These studies suggest that the underprivileged are in higher danger of motor vehicle accidents, however geographically we may see this being more prevalent in areas with a sharp gradient from low income to high.

2.2.5 Education

While strong correlations can be made with education and traffic crashes, other demographic factors may influence these results more than previously observed. Prior research has found that for females there is no observed correlation between education level and crash risk, while for males, lower education rates are correlated to higher crash risks (Sami, et al, 2013). However, it has already been established that young males are the most at risk, and the correlation is similar between education and age as younger people tend to have a lower education level. With no correlation for crash rates or risk when looking at females age and education level, this strongly suggests serial autocorrelation between education and age.

2.2.6 Driver’s License Possession

Research conducted in Japan found that pedestrians within the 65-74 age group (the only age group studied) were more likely to be victims of traffic crashes if they were not holders of a valid driver’s license (Retting, Ferguson and McCartt, 2011). Those who did not have prior knowledge of using a car had less empathy with drivers and had less understanding of “how others perceive oneself and what they intend to do next,” therefore they found themselves involved in more crashes.

While there is a depth of research examining demographics and motor vehicle crashes, nothing comprehensively looks at all categories to determine what is most likely to predict crashes. Despite this, a common theme tends to emerge when looking at the previous research and two factors emerge more than others. Gender and age seem to be the primary factors which determine the crash risk and rate with both young males and the elderly being most at risk. Despite not being the primary focus of the articles in this review, many suggested age and sex were the best predictors of crash risk and rate.

2.3 Study Methods

This research employed several statistical analysis methods, including summary statistics and binary regression models, to describe trends in the data as well as make predictions regarding correlation and causality between variables. Each method is described in detail below and was selected based on its appropriateness for use with study-specific data and the research questions and hypotheses.

2.3.1 Summary Statistics

Summary statistics are used to provide a quick and simple description of the data without any predictive component or significance testing. They may include mean (average), median (center point of data), mode (most frequently occurring value), minimum value, maximum value, value range, standard deviation, and frequency percentages. Summary statistics were used in this analysis to provide context for the crash data, and demographics.

2.3.2 Pearson's Chi-Square Test

A Chi-Square test is used on categorical data to compare an observed distribution to a theoretical one (measuring goodness of fit) for one or more categories. The events included must be mutually exclusive (e.g., weather cannot be clear and raining at the same time) and have a total probability of 1 (Greene, 2015).

Model:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

where

χ^2 is the chi-square value

Σ is the summation sign

O is the observed frequency

E is the expected frequency

2.3.3 Binary Logistic Regression

Binary logistic regression is used to estimate the odds or probability that a characteristic is present given the values of explanatory variables (Greene, 2015). In this research, the probability of a specific type of crash occurring will be predicted based on the presence of specific demographic characteristics (e.g. age groups, ethnic groups, educations levels, etc.). The statistical model is derived as follows:

Variables:

Y_i = 1 if a specific type of crash (*i*) occured
Y_i = 0 if a specific type of crash (*i)* did not occur

$X = (X_1, X_2, ..., X_k)$ will be a set of explanatory variables which can be discrete, continuous, or a combination (outlined in Table 1). x_i is the observed value of the explanatory variables for observation *i*.

Model:

$$\pi_i = Pr(Y_i = 1|X_i = x_i) = \frac{exp(\beta_0 + \beta_1 x_i)}{1 + exp(\beta_0 + \beta_1 x_i)}$$

or,

$$logit(\pi_i) = log\left(\frac{\pi_i}{1 - \pi_i}\right)$$

$$= \beta_0 + \beta_1 x_i$$

$$= \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k + \beta_k x_{ik}$$

Assumptions:

- The data $Y_1, Y_2, ..., Y_n$ are independently distributed (cases are independent)
- Distribution of Y_i is $Bin(n_i, \pi_i)$, i.e., binary logistic regression model assumes binomial distribution of the response. The dependent variable does NOT need to be normally

distributed, but it typically assumes a distribution from an exponential family (e.g. binomial, Poisson, multinomial, normal, etc.)

- Does NOT assume a linear relationship between the dependent variable and the independent variables, but it does assume linear relationship between the logit of the response and the explanatory variables; $logit(\pi) = \beta_0 + \beta X$
- Independent (explanatory) variables can even be the power terms or some other nonlinear transformations of the original independent variables.
- The homogeneity of variance does NOT need to be satisfied. In fact, it is not even possible in many cases given the model structure.
- Errors need to be independent but NOT normally distributed.
- It uses maximum likelihood estimation (MLE) rather than ordinary least squares (OLS) to estimate the parameters, and thus relies on large-sample approximations.
- Goodness-of-fit measures rely on sufficiently large samples, where a heuristic rule is that not more than 20% of the expected cells counts are less than 5 (Greene, 2015).

2.4 Summary

Existing research shows that males are disproportionately involved in more vehicle crashes, with WHO reporting around three quarters (73%) of fatalities being male; although females commit more unintentional errors in their driving. Traffic crashes are the leading cause of death for individuals between the age of 5-29, and as drivers gain more experience on a road (with age) their performance improves.

Fatalities are highest among Native American drivers as a percentage of all deaths, independent from socioeconomic factors. Fatally injured drivers who had been drinking is also highest among Native Americans, who were also found to be less likely to hold a valid US driver's license or wear a seatbelt. Hispanic populations have the second highest rates of fatalities, fatal crashes while driving under the influence, and seatbelt non-use. Research suggests this may be due to confusion in signal and signage understanding and cultural traditions of non-compliance with roadway signals and signage.

Prior research has found no observed correlation between education level and crash risk; however, the literature shows that lower income individuals experience more severe injuries and fatalities when involved in crashes, likely due to the types of vehicles involved.

This research employs several statistical analysis methods to describe trends in the data as well as make predictions regarding correlation and causality between variables. Each method was selected based on its appropriateness for study-specific data and the research questions and hypotheses. Methods used in this research include descriptive statistics, chi-square analysis, and binary logistic regression models.

3.0 DATA COLLECTION

3.1 Overview

This chapter discusses the data collected for the research and presents an overview of descriptive characteristics for the study area and a discussion of data quality. The overview includes a description of the geographic scale of the data collection, a summary of the demographics data used, and a description of the crash data and covariates from that dataset included in the subsequent analysis.

3.2 Study Site Identification

This research examines demographics and crash risk in Salt Lake County, Utah (Figure 1). Salt Lake County is located between the Wasatch and Oquirrh Mountains, and encompasses the core of Utah's urban population. Utah's capital city, Salt Lake City, is located on the north end of Salt Lake County.

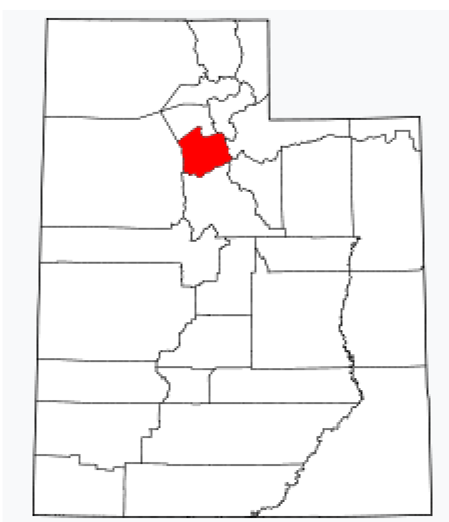

Figure 1. Salt Lake County

Communities in Salt Lake County exhibit a wide range of demographic characteristics. The Salt Lake City School District reports over 80 languages spoken in the homes of its students

(Gardner Policy Institute, 2017). This depth and breadth of cultural, linguistic, ethnic, and intellectual diversity are unprecedented in Utah, making it the ideal study site for this research.

3.3 Demographic Data

All demographic data was provided by the Wasatch Front Regional Council (WFRC) and the Kem C. Gardner Policy Institute at the University of Utah. Data was derived from the 2010 U.S. Census and American Community Survey (ACS). The ACS provides yearly information and estimates between census years regarding job and occupations, educational attainment, home ownership, as well as other characteristics. Four modes are used in ACS data collection: 1. Internet 2. Mailout/Mailback 3. Computer Assisted Telephone Interview (CATI), and 4. Computer Assisted Personal Interview (CAPI) (U.S. Census, 2017). The ACS employs a dual-phase, dual-stage sample design. The first-phase creates a sample based on two separate address groups at different points in time. Both samples are selected in two stages of sampling, a first-stage and a second-stage. Following the second-stage sampling, most of the sample addresses are randomly assigned to one of twelve months in the sample year (addresses in rural Alaska are assigned to either January or September). The second-phase of sampling occurs when the CAPI sample is selected (U.S. Census, 2017). Weights are applied to the sample based on each sample person and each sample housing unit. Estimates of person characteristics are based on a person weight. Estimates of family, household, and housing unit characteristics are based on a housing unit weight.

Data used in this evaluation focused on Salt Lake County and included both year to year estimates and 5-year projections. The data used in this study is current as of 2018. All census data is geographically referenced to the census tract level but can be aggregated to all larger geographic scales. The demographic data was geo-referenced down to the census tract level for the highest allowable accuracy.

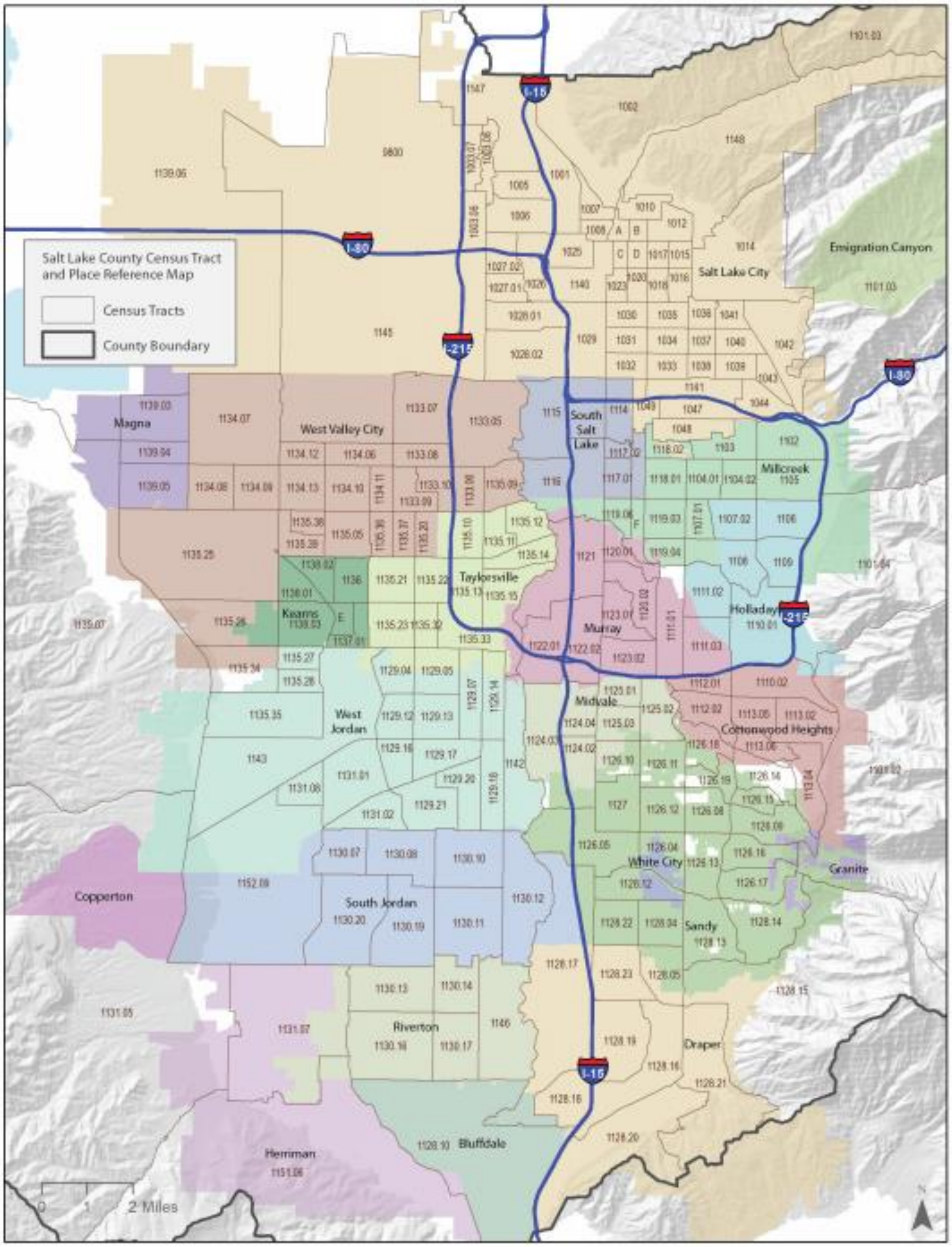

Figure 2. Salt Lake County Census Tract and Place Reference Map

Salt Lake County is made up of 212 census tracts (shown in Figure 2). This spatial stratification allows for identifying detailed small-scale household information at nearly a neighborhood level. Census tracts are similar in size, as each contains a similar number of households. Aerial size may vary quite a bit in less populated areas (Gardner Policy Institute, 2017).

The next source of data used in this analysis is the socio-economic (SE) current and projected data from the Wasatch Front Regional Council's (WFRC) regional travel demand model. This dataset is identified and projected at the spatial scale of Traffic Analysis Zones (TAZs). Data collected at the TAZ level was merged with the census tract level data for each crash. For the most part, the TAZs line up directly with the boundaries of the census tracts. In cases where boundaries did not line up exactly, or where an overlap occurred, the TAZ was converted into a point file and a corresponding census tract was identified based on the centroid of the TAZ point. A list of demographic variables collected for this research is shown in Table 1 below.

Table 1. Study Demographic Characteristics

Population Characteristics	**Household Characteristics**	**Other Characteristics**
Age	Household size	Unemployment
Children-% Under 14	Household Income	Working multiple jobs
Teens- % Age 15-19	Households living in Poverty	Journey to Work (mode)
Young Adults- % 20-34	Marital Status	Auto
Middle Adults- % 35-49		Single Occupancy
Late Adults- % 50-64		Carpool
Seniors- % Age 65+		Transit
Sex (% male)		Walked
Race		Biked
White		Employment Sector
Black		Blue collar
Native American		White collar
Polynesian		Home based
Asian		
Other		
Multiple races		
Hispanic or Latino		
Educational Attainment		
Currently enrolled in middle and high school		
Less than high School		
High school graduate		
Some college		
Associates degree		
Bachelor's degree		
Graduate or Professional School		

3.4 Crash Data

Vehicle crash data was identified using UDOT's SafeMap tool – a comprehensive data analytics system that stores and allows queries of statewide crash data. All crashes occurring within Salt Lake County between January 1, 2010 and December 31, 2018 were flagged and tallied. A total of 215,645 crashes were identified. Because some crashes occurred on the boundary of 2 census tracts, they were included in both census tracts for the statistical analysis so as not to eliminate the role the demographics from either tract may play. In total 305,607 crashes were included in the statistical analysis to maximize the potential to identify correlations between both crash and local area characteristics.

3.5 Data Quality

All data included in this analysis was collected outside the scope of this research. The data was collected by professional organizations and has been cleaned and verified for accuracy and validity. For example, the American Community Survey results are confirmed with a margin of error within a 90% confidence interval (Berkley, 2017). ACS produces estimates of the actual figures that would be obtained by interviewing the entire population. The estimates are the result of measuring the sample and are subject to variation. Sampling error in data arises due to the use of probability sampling, which is necessary to ensure the integrity and representativeness of sample survey results. According to the U.S. Census Bureau, they "protect against the effect of systematic errors on survey estimates by conducting extensive research and evaluation programs on sampling techniques, questionnaire design, and data collection and processing procedures (U.S. Census, 2017)."

The primary data source for the FARS database is the police crash report. Analysts also collect additional documentation such as vehicle registration, driver history, and vital statistics data on each FARS case from several state agencies. For FARS data to be most useful, it is critical that the quality of data is maintained. However, NHTSA (McDonough and Smith, 2010) has noted that the size of the collection effort and the geographical disbursement of the FARS analysts make this task difficult. To help ensure the quality of the data, analyst training, documentation of system standards, and data monitoring is conducted annually. Recent appropriations have provided NHTSA with funding to improve FARS data quality by enhancing and updating its quality control processes. This FARS case re-coding process has enabled NHTSA to check the accuracy of the original coding, assess analyst performance, and conduct state specific training to address problems identified in the sampling (McDonough and Smith, 2010).

Utah crash data analyzed in this report are taken from the Utah crash data network compiled by UDOT and UDPS. This data is gathered and imported from crash reports completed at the scene of each crash by law enforcement officers (LEOs). Prior research has shown several threats to the validity and reliability of that data, particularly for crashes involving non-motorists. For example, it was shown that "only 33% of the crash reports analyzed contained no errors, and 15% of crash reports involving a pedestrian fatality contained three or more coding errors (Burbidge, 2016). Additionally, the categories on the crash form provide many options to choose from which can in fact reduce the accuracy of the reporting. In the category "first harmful event" there are nearly 70 options for the LEO to choose from that range from circumstances that could occur (e.g. downhill runaway, crossed median/centerline, equipment failure) to objects involved in a collision (e.g. animal, parked motor vehicle, guardrail, etc). This category encompasses so many options that are not mutually exclusive, that it can be difficult to identify just one on the crash report.

As any analysis is only as good as the data used, it is critical to understand the limitations and drawbacks present in the dataset. While there are several potential drawbacks within this dataset, these sources have been widely used as industry standard and are not typically questioned in terms of their accuracy or reliability. Where questions arise in the analysis results that may be attributed to these limitations, they are identified and discussed.

3.6 Summary

This research examines demographics and crash risk in Salt Lake County, Utah. Data sources used in this analysis include demographic data from the U.S. Census and the American Community Survey. Additional Socio-economic data was compiled from projections in the Wasatch Front Regional Council's travel demand model. Vehicle crash data was identified using UDOT's SafeMap tool (Numetric), for all crashes occurring between January 1, 2010 and December 31, 2018. Although there are several limiting factors for the included datasets, all data used in this study is accepted as industry standard and are not typically questioned.

4.0 DATA EVALUATION

4.1 Overview

This section includes analysis of all demographic and crash data. First, descriptive statistics are provided describing the crash data in the sample. Next, statistical methods are used to identify significant correlations between demographic characteristics and crash types.

4.2 Summary of Crash Data

Between 2010-2018 over 215,000 crashes occurred. Of those crashes, a large majority resulted in no easily identifiable injury to the vehicle occupants or non-motorists (87.4%). As shown in Table 2 below, approximately one in ten people experienced a minor injury (10.6%), with 2% of crashes resulting in serious injury (1.7%), or fatality (0.3%).

Table 2. Summary of Crash Characteristics

Characteristic	**Percentage**
Crash Severity	
No Injury	68.4
Possible Injury	19.0
Minor Injury	10.6
Serious Injury	1.7
Fatal	0.3
N=215,645	

4.2.1 Crash Severity

Because a large percentage of total crashes result in minor injuries or less, it is important to determine which crash types are more likely to result in serious injury or death. By determining which crash types are most likely to result in severe outcomes, the subsequent analysis can identify which demographic characteristics are correlated and concomitantly where these crashes may be most likely to occur. Table 3 below shows the breakdown of crash types within the sample. The first column provides the percentage of total crashes, while the second column shows the percentage resulting in severe injuries, and column three shows the percentage of fatal crashes. Nearly a third of all fatal crashes involve alcohol (27.4%). This is notable because alcohol was only recorded in 1.2% of all crashes that occurred. Other overly dangerous crash types include collision with a fixed object, crashes where an older driver or teen driver is involved, crashes involving unrestrained driver/passengers, and those where a traffic control device was disregarded.

Table 3. Crash Types - by Severity

Crash Type	% Total Crashes	% Serious Injuries	% Fatal Crashes
Aggressive Driving	1.2	3.5	5.8
Alcohol Crash	3.6	9.4	27.4
Older Driver Involved	12.5	14.4	17.1
Teen Driver Involved	19.4	16.3	13.4
Collision with a Fixed Object	11.7	13.4	20.2
Disregard Traffic Control Device	7.8	16.3	13.0
Distracted Driving	9.4	10.5	8.2
Drowsy Driving	1.3	1.4	1.5
Improper Restraint	1.1	4.1	7.1
Unrestrained	2.6	9.0	16.5
Total Crashes	215,645	5,324	777

There were several other notable characteristics of serious and fatal crashes that are likely unrelated to demographics, but nevertheless provide context for the circumstances surrounding these highly dangerous crashes. Over half of all fatal collisions involve a single vehicle (52%) and 25.7% of serious crashes involve a single vehicle, while single vehicle crashes make up only 17% of total crashes in the sample. Also, over 20% of fatal crashes occurred at night in dark lighting conditions.

4.2.2 Speed

Utah’s crash reporting paradigm does not include speed under the aggressive driving definition. Therefore, it is important to note if speed plays an integral role in crash severity. According to the sample data, the mean estimated speed at the time of the crash (as calculated by law enforcement) was 19.78mph (32 kph). The macro scale of this data is not entirely useful, as travel speed and roadway speed limit can vary widely from one location to another. Therefore, it is vital to determine how fast a vehicle is traveling relative to the roadway and environment. For the purpose of this analysis, relative speed is calculated by dividing the speed at the time of the crash (S_{crash}) by the posted speed limit (psl), in miles per hour, on the roadway where the crash occurs (S_{psl}).

$$Relative\ Speed = \frac{S_{crash}}{S_{psl}}$$

If a driver was traveling faster than the speed limit at the time of the crash, the relative speed would be greater than 1. The mean relative speed for all crashes was 0.45, meaning that vehicles were traveling at 45% of the posted speed limit at the time of the crash. This is consistent with prior research showing that most vehicles are slowing down at the time of a crash rather than traveling a consistent speed (Burbidge, 2016). It is also consistent with the crash data which shows that front to end collisions are the most frequently occurring type (Table 4). This is expected, as vehicles typically slow as they approach the car in front of them.

4.2.3 Manner of Collision

Next, the manner of collision was evaluated, based on crash severity. The most severe injuries occurred in crashes involving an angle impact or a single vehicle crash. Approximately 36% of all serious and fatal crashes involve an angle impact, while almost 38% involve a single vehicle (when serious and fatal crashes are pooled).

Table 4. Manner of Collision – by Severity

Crash Type	% Total Crashes	% Serious Injuries	% Fatal Crashes
Angle	26.6	35.3	24.6
Front to Rear	36.8	15.1	7.5
Head On (front to front)	2.3	6.5	9.0
Parked Vehicle	4.9	3.3	4.9
Rear to Rear	0.1	0.0	0.0
Rear to Side	0.2	0.1	0.0
Sideswipe Opposite Direction	1.3	1.1	0.6
Sideswipe Same Direction	9.9	2.9	1.4
Single Vehicle	18.0	35.7	52.0
Total Crashes	215,645	5,324	777

Most safety programs and infrastructure interventions focus on improving safety by reducing or eliminating crashes between vehicles. This data shows that a majority of fatal crashes involve a single vehicle. Pairing this information with the crash types paints a broader picture of the circumstances surrounding serious and fatal crashes.

The following section will investigate these characteristics relative to the environments in which these crashes occur. Statistical models will be used to correlate neighborhood characteristics with crash types and identify where these types of crashes are most likely to occur.

4.3 Demographic Data

For all statistical models, demographic variables were included as predictors evaluated for correlation with specific crash characteristics. The demographics included were chosen based on preliminary testing which identified the most likely predictors of variation within the crash data. Additional variables were included in earlier iterations of the models (see Table 1) but were not included in the final evaluating due to lack of statistical significance. Each characteristic is provided as a percentage of the population in the census tract where the crash occurred. However, for several of the models, age was converted to a loglinear elasticity showing the percentage for that group was low, medium, or high, as it better represented the data for that model. The following demographic characteristics were included:

- Sex (percent male)
- Age
  - Under 14
  - Teens (15-19)
  - Young Adults (20-34)
  - Middle Adults (35-49)
  - Late Adults (50-64)
  - Seniors (65+)
- Educational Attainment
  - Less than High School
  - High School Graduate
  - Some College
  - Associates Degree
  - Bachelor’s Degree
  - Graduate or Professional School
- Unemployed
  - % of population
- Ethnicity
  - White
  - Black

- Native American
- Polynesian
- Asian
- Other
- Multiple
- Hispanic or Latino

- Household Size
  - # of persons per household

- Commute Mode
  - Single Occupant Vehicle (SOV)
  - Carpool
  - Transit
  - Walk
  - Bike

Employment sector, disability, and household income were included in preliminary runs of all models, but the variables were not significant and therefore are not reported in the analysis tables shown in the following sections. Also, the percentage of females was not included in the models as there is a direct inverse correlation to the percentage of males.

4.4 Demographics Crash Type

First, a thorough analysis of local demographic characteristics and crash types was conducted. Based on the preliminary crash data, the following section focuses on DUI crashes, crashes involving older drivers, crashes which involve a collision with a fixed object, and unrestrained crashes, as these are the most likely to include a serious or fatal outcome.

4.4.1 Driving Under the Influence (DUI)

A binary logistic regression was performed to predict the probability of a crash involving DUI based on the characteristics of those living nearby. The model was statistically significant and correctly classified over 96.4% of cases ($X^2 = 28.032$).

Table 5. Demographics and DUI Crashes

Demographics	B	Sig.	Exp(B)	95% Wald Confidence Interval	
				Lower	Upper
Percent Male	-1.014	0.006	0.363	0.176	0.747
Age					
Under14	-2.393	0.132	0.091	0.004	2.054
Teens	-3.267	0.044	0.038	0.002	0.920
Young Adults	-1.217	0.435	0.296	0.014	6.291
Middle Adults	-1.298	0.400	0.273	0.013	5.619
Late Adults	-1.884	0.224	0.152	0.007	3.165
Seniors	-1.613	0.309	0.199	0.009	4.468
Educational Attainment					
Less than High School	2.075	0.056	7.962	0.946	67.033
High School Graduate	1.870	0.085	6.488	0.774	54.386
Some College	-0.371	0.724	0.690	0.088	5.430
Associates Degree	0.382	0.759	1.465	0.127	16.862
Bachelor’s Degree	3.113	0.004	22.479	2.639	191.442
Graduate or Professional	1.774	0.114	5.893	0.653	53.152
Unemployed	0.984	0.070	2.674	0.836	8.549
Ethnicity					
White	-2.363	0.066	0.094	0.008	1.173
Black	-2.131	0.155	0.119	0.006	2.246
Native	0.266	0.854	1.304	0.077	21.949
Polynesian	-1.661	0.222	0.190	0.013	2.731
Asian	-2.247	0.092	0.106	0.008	1.440
Other	-0.663	0.609	0.515	0.041	6.548
Multiple	-1.431	0.297	0.239	0.016	3.526
Hispanic	-0.888	0.000	0.412	0.274	0.619
Household Size	-0.192	0.000	0.825	0.789	0.863
Commute Mode					
Single-Occupant Vehicle	0.770	0.718	2.160	0.033	141.894
Carpool	0.816	0.704	2.260	0.034	151.852
Transit	-0.758	0.257	0.469	0.126	1.738
Walked	0.222	0.737	1.249	0.341	4.574
Bike	4.169	0.000	64.624	7.977	523.523
Constant	12.018	0.000	165,780.62		

When interpreting the models in this section, each table provides the likelihood ratio that a crash will occur in a given area (*B*). A positive number means that as the percentage of the demographic increases, it is more likely that the given crash type will occur, a negative number means it is less likely. Significance is shown as a decimal with anything smaller than 0.05 being statistically significant (within a 95% confidence interval). Exp(*B*) shows the log evaluation of the B value (for statistical purposes). Lastly, the confidence interval is shown which encompasses 95% of all values for the variable in question. This means that 95% of the observations for the demographic fall within the lower and upper bounds these numbers show.

As shown in Table 5, as the percentage of teen residents and Hispanic residents in a census tract increases, the probability of DUI crashes within the area significantly decreases. Also, as the average household size increases, the probability of DUI crashes occurring in the area decreases. Transportation modes and educational attainment also play a role. As the percentage of cycle commuters and individuals with a bachelor's degree increases, the probability of DUI crashes occurring in the same census tract increases.

4.4.2 Older Drivers

Next, a binary logit regression was run to evaluate the relationship between demographics and crashes involving older drivers. The model was statistically significant (X^2 = 76.712) and correctly classified over 87.5% of cases. Table 6 shows that as the percentage of men in a census tract increases, the probability of crashes involving older drivers significantly decreases. Alternatively, for each percentage increase in the young adult (ages 20-34) and senior populations (age 65+), the probability of crashes involving older drivers increases significantly.

Educational attainment was also correlated to crashes involving older drivers. As the percentages of residents with less than a graduate degree increased, the probability of crashes involving older individuals increased. Areas with higher unemployment and larger Native America, Polynesian, Asian and Hispanic populations (as well as those identifying multiple races or other races) are correlated to a lower probability of crashes involving older individuals. Census tracts where alternative transportation modes (carpool, transit, walking, biking) are used at a higher rate, are significantly correlated with a lower probability of crashes involving older drivers as well.

Table 6. Demographics and Older Driver Crashes

Demographics	B	Sig.	Exp(B)	95% Wald Confidence Interval	
				Lower	Upper
Percent Male	-1.314	0.000	0.269	0.174	0.414
Age					
Under14	1.267	0.154	3.552	0.622	20.268
Teens	1.538	0.090	4.657	0.786	27.587
Young Adults	2.214	0.011	9.154	1.655	50.644
Middle Adults	1.523	0.079	4.585	0.839	25.064
Late Adults	1.103	0.203	3.013	0.551	16.479
Seniors	6.209	0.000	496.976	87.654	2817.740
Educational Attainment					
Less than High School	0.117	0.468	1.124	0.820	1.539
High School Graduate	0.382	0.013	1.466	1.084	1.980
Some College	2.067	0.000	7.902	6.119	10.205
Associates Degree	1.032	0.004	2.808	1.392	5.665
Bachelor's Degree	2.596	0.000	13.407	9.604	18.715
Graduate or Professional	-0.086	0.676	0.918	0.615	1.371
Unemployed	-0.928	0.007	0.395	0.200	0.780
Ethnicity					
White	-1.182	0.112	0.307	0.071	1.318
Black	-0.172	0.845	0.842	0.151	4.693
Native	-2.199	0.010	0.111	0.021	0.587
Polynesian	-1.873	0.020	0.154	0.032	0.742
Asian	-2.359	0.002	0.094	0.021	0.428
Other	-2.357	0.002	0.095	0.022	0.410
Multiple	-2.312	0.004	0.099	0.021	0.472
Hispanic	-0.321	0.006	0.726	0.576	0.914
Household Size	0.024	.101	1.024	0.995	1.054
Commute Mode					
Single-Occupant Vehicle	-2.363	0.056	0.094	0.008	1.061
Carpool	-2.839	0.022	0.058	0.005	0.669
Transit	-1.535	0.000	0.215	0.103	0.450
Walked	-3.830	0.000	0.022	0.010	0.047
Bike	-2.216	0.001	0.109	0.028	0.422
Constant	-6.617	0.00	0.001		

4.4.3 Collision with a Fixed Object

Once again, a binary logit regression was employed to evaluate the relationship between demographics and collisions with a fixed object. The model was statistically significant (X^2 = 297.371) and correctly classified over 88.5% of cases.

Table 7. Demographics and Collision with a Fixed Object

Demographics	B	Sig.	Exp(B)	95% Wald Confidence Interval	
				Lower	Upper
Percent Male	3.039	0.000	20.876	13.465	32.365
Age					
Under14	0.437	0.627	1.549	.265	9.054
Teens	-2.063	0.026	0.127	.021	.783
Young Adults	-0.977	0.270	0.376	.066	2.139
Middle Adults	-0.550	0.534	0.577	.102	3.266
Late Adults	2.644	0.003	14.065	2.477	79.866
Seniors	-1.186	0.188	0.306	.052	1.788
Educational Attainment					
Less than High School	2.043	0.001	7.715	2.224	26.759
High School Graduate	1.858	0.003	6.410	1.853	22.180
Some College	1.146	0.062	3.146	0.945	10.469
Associates Degree	1.356	0.062	3.881	0.932	16.159
Bachelor's Degree	2.700	0.000	14.883	4.282	51.735
Graduate or Professional	6.346	0.000	570.418	157.433	2066.770
Unemployed	3.809	0.000	45.087	22.614	89.892
Ethnicity					
White	1.894	0.013	6.649	1.500	29.461
Black	-0.971	0.275	.379	.066	2.167
Native	4.637	0.000	103.216	19.608	543.324
Polynesian	4.571	0.000	96.671	20.110	464.704
Asian	0.749	0.342	2.116	0.451	9.918
Other	5.206	0.000	182.428	40.773	816.232
Multiple	3.461	0.000	31.845	6.478	156.546
Hispanic	-2.388	0.000	0.092	0.072	0.117
Household Size	-0.473	0.000	0.623	0.607	0.639
Commute Mode					

Single-Occupant Vehicle	13.830	.000	1014921.54	80254.926	12834922.32
Carpool	15.956	.000	8500776.63	657296.33	109940067.59
Transit	-.228	.577	.796	.358	1.772
Walked	-3.076	.000	.046	.021	.103
Bike	-1.067	.129	.344	.087	1.366
Constant	6.420	.001	613.895		

Each percentage increase in the male population results in over a 300% increase in the probability of crashes involving collision with a fixed object. Likewise, areas with a higher percentage of the population of teenagers and late adults (between the ages of 50-64) have a significantly higher probability of having crashes involving collision with a fixed object.

The effects of education were rather dichotomous. Areas with low educational attainment (high percentage of the population with high school or less) and areas with very high educational attainment (bachelor's degree or higher), were both significantly correlated to a higher probability of collisions involving a fixed object. Additionally, for each percentage that the unemployment rate increased, there was a 4% increase in the probability of a collision with a fixed object occurring in the census tract.

An evaluation of ethnicity shows that as the percentage of Hispanics increases, the probability of these types of crashes decreases. While an increase in the population of whites, Native Americans, Polynesians, and Multiple/Other races increases, so does the likelihood of this crash type. Additionally, as the average number of people per household increases, the probability of these types of crashes significantly decreases.

Employment status was negatively correlated, suggesting that as the percentage of the population in an area who is unemployed increases, the probability of these types of crashes decreases. This is likely tied to the auto ownership and usage, as the analysis revealed that census tracts with higher rates of auto commuting (driving alone or carpooling) were significantly correlated to a dramatic increase (>1300%) in the probability of a crash involving collision with a fixed object occurring in the same area. As the percentage of walking commuters increased the probability of this type of collision occurring nearby decreased by over 300%.

4.4.4 Unrestrained Crashes

Next, a binary logit regression was run to evaluate the relationship between demographics and the occurrence of unrestrained crashes. The model was statistically significant (X^2 = 30.522) and correctly classified over 97.4% of cases. It should be noted that this model was calibrated differently than the models described above due to the age distribution in the unrestrained crash category. In order to isolate and refine the potential impact of age groups and control for endogeneity, age groups for each census tract were transformed using an elasticity method to classify each as low, medium, or high in terms of representation by a given age group. For example, a census tract would be classified as follows for youth (under age 14): Low = $\leq$ 19%, Medium = 20-26%, High = $\geq$27%. Classification groups were configure based on quartiles within the sample distribution of each age group, therefore percentages and ranges are not the same for each age group. However, each measurement is relative to the representation shown in Salt Lake County.

Census tracts with a higher percentage of middle adults (ages 35-49) had a significantly lower probability of unrestrained crashes, while tracts with higher percentages of children and seniors had a significantly higher probability of unrestrained crashes.

As the percentage of Native Americans increased in a census tract there was a significant increase in the probability of unrestrained crashes in the area. Alternatively, as the percentage of non-motorized commuters increased (walking/biking), the probability of unrestrained crashes increased significantly.

Table 8. Demographics and Unrestrained Crashes

Demographics	B	Sig.	Exp(B)	95% Confidence Interval	
				Lower	Upper
Percent Male	-0.197	0.638	0.821	0.361	1.867
Age					
Under14	0.125	0.000	1.133	1.096	1.172
Teens	0.013	0.408	1.013	0.982	1.044
Young Adults	-0.023	0.248	0.978	0.941	1.016
Middle Adults	-0.051	0.002	0.950	0.919	0.982
Late Adults	-0.010	0.571	0.990	0.956	1.025
Seniors	0.050	0.008	1.051	1.013	1.091
Educational Attainment					
Less than High School	-.283	.817	.754	.068	8.293
High School Graduate	1.690	.166	5.420	.497	59.104
Some College	-1.994	.090	.136	.014	1.368
Associates Degree	-.887	.535	.412	.025	6.796
Bachelor's Degree	.492	.689	1.636	.147	18.267
Graduate or Professional	.583	.643	1.791	.152	21.113
Unemployed	.531	.441	1.700	.440	6.564
Ethnicity					
White	-.780	.627	.458	.020	10.640
Black	-.727	.702	.483	.012	20.022
Native	3.695	.049	40.232	1.011	1601.233
Polynesian	-.910	.597	.403	.014	11.689
Asian	-1.521	.362	.218	.008	5.771
Other	-.559	.731	.572	.023	13.933
Multiple	-1.589	.378	.204	.006	6.969
Hispanic	-.206	.531	.814	.428	1.550
Household Size	-0.068	0.007	0.934	0.88	0.981
Commute Mode					
Single-Occupant Vehicle	1.244	.615	3.468	.027	438.523
Carpool	.086	.972	1.090	.008	144.013
Transit	.876	.241	2.402	.555	10.400
Walked	4.124	.000	61.789	13.618	280.351
Bike	10.432	.000	33923.742	3022.632	380734.498
Constant	8.003	.003	2989.689		

4.4.5 Disregarding a Traffic Control Device

Lastly, a binary logit regression was run to evaluate the relationship between demographics and the crashes where a traffic control device was disregarded. The model was statistically significant ($X^2 = 213.77$) and correctly classified over 92.2% of cases.

Areas with a larger percentage of males and young adults (ages 19-24) were less likely to have crashes involving disregard for traffic devices, while census tracts with a higher percentage of children, and those 25 and older were correlated to a higher probability of these crashes.

Looking at education we find that areas with a higher percentage of residents with less than an associate's degree, or a high percentage with a bachelor's Degree had a higher probability of crashes disregarding traffic control devices. Tracts with a higher percentage of the population holding graduate degrees or with higher unemployment had a lower probability of these types of crashes.

Table 9. Demographics and Disregarding Traffic Control Devices

Demographics	B	Sig.	Exp(B)	95% Confidence Interval	
				Lower	Upper
Percent Male	-1.173	0.000	0.309	0.188	0.510
Age					
Under14	0.080	0.000	1.084	1.060	1.108
Teens	0.012	0.296	1.012	0.990	1.034
Young Adults	-0.024	0.123	0.976	0.946	1.007
Middle Adults	0.078	0.000	1.081	1.058	1.104
Late Adults	0.046	0.001	1.047	1.020	1.075
Seniors	0.100	0.000	1.105	1.073	1.137
Educational Attainment					
Less than High School	1.468	0.000	4.339	2.096	8.985
High School Graduate	2.055	0.000	7.803	4.404	13.827
Some College	0.979	0.001	2.661	1.513	4.680
Associates Degree	-0.368	0.513	0.692	0.230	2.085
Bachelor's Degree	2.511	0.000	12.311	6.641	22.822
Graduate or Professional	-1.670	0.000	0.188	0.093	0.381
Unemployed	-3.018	0.000	0.049	0.022	0.109
Ethnicity					

White	-10.033	0.000	0.000	0.000	0.000
Black	-11.470	0.000	0.000	0.000	0.000
Native	-2.549	0.024	0.078	0.009	0.715
Polynesian	-8.560	0.000	0.000	0.000	0.001
Asian	-9.468	0.000	0.000	0.000	0.001
Other	-9.733	0.000	0.000	0.000	0.000
Multiple	-8.595	0.000	0.000	0.000	0.002
Hispanic	-1.157	0.000	0.315	0.211	0.469
Household Size	0.154	0.000	1.166	1.128	1.205
Commute Mode					
Single-Occupant Vehicle	0.339	0.262	1.403	.776	2.538
Carpool	0.221	0.527	1.247	.629	2.475
Transit	2.679	0.000	14.575	6.221	34.148
Walked	2.492	0.000	12.083	5.142	28.390
Bike	7.148	0.000	1271.219	298.061	5,421.703
Constant	5.940	0.000	379.877		

All ethnicities were correlated to a reduction in the probability of crashes involving a disregard for traffic control however with varying degree of impact. Hispanic and Native American populations had the lowest impact, while Whites and Blacks had the highest impact. Larger household size was correlated to an increase probability of these crash types.

Lastly, journey to work data shows that areas with higher rates of transit use or walking and biking are correlated to a significant increase in the probability of crashes involving a disregard of traffic signals. Each percentage increase in bicycle commuting results in a 700% increase in probability of these crashes occurring nearby.

4.5 Demographics and Manner of Collision

The next goal of this research was to identify the relationship between area demographics and the manner of collisions that occur in the area. To this end, several binomial regression models were employed to identify which characteristics may impact the probability of angle crashes or single vehicle crashes occurring, as these are the most dangerous types of collisions.

4.5.1 Angle Crash

An angle crash occurs when two vehicles arrive on perpendicular roads and collide. There are two main types of angle crashes (Figure 3 below); one where entering traffic has stopped, and one where entering traffic disregards a stop or signal. Angle crashes were specifically included in this evaluation due to their overrepresentation in serious and fatal crashes (shown in Table 4).

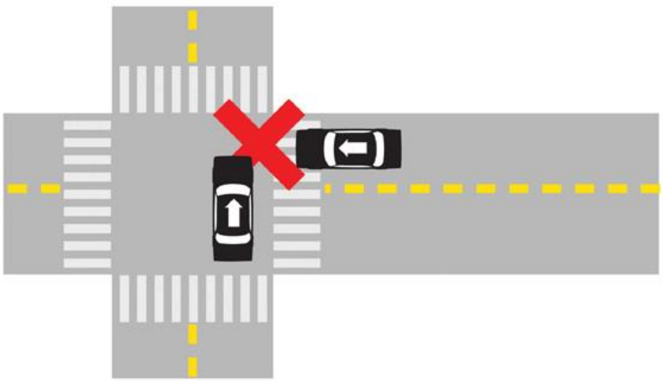

Figure 3. Angle Crash

A binary logit regression was run to evaluate the relationship between demographics and unrestrained crashes (See Table 10). The model was statistically significant (X^2 = 814.70) and correctly classified over 99.5% of cases. This model once again employed age group categories (low, medium, high) as an elasticity in the model to control for skewness and improve mode predictability. As shown in Table 10, an increase in the representation of each group, except for young adults) was significantly correlated with a higher probability of angle crashes occurring nearby. However, the coefficients are small, suggesting marginal effects which can be attributed to the incredibly large sample size (N=305,607).

Table 10. Demographics and Angle Crashes

Demographics	B	Sig.	Exp(B)	95% Confidence Interval	
				Lower	Upper
Percent Male	-1.364	0.000	0.256	0.188	0.347
Age					
Under14	0.046	0.000	1.047	1.034	1.059
Teens	0.041	0.000	1.042	1.031	1.054
Young Adults	-0.001	0.908	.999	.985	1.013
Middle Adults	0.018	0.003	1.019	1.006	1.031
Late Adults	0.021	0.001	1.022	1.009	1.035
Percent Seniors	0.147	0.000	1.158	1.142	1.174
Educational Attainment					
Less than High School	-3.014	0.000	0.049	0.021	0.117
High School Graduate	1.253	0.005	3.500	1.463	8.372
Some College	.591	0.166	1.805	0.783	4.160
Associates Degree	-1.794	0.001	0.166	0.060	0.458
Bachelor’s Degree	.055	0.902	1.057	0.439	2.547
Graduate or Professional	-3.880	0.000	0.021	0.008	0.051
Unemployed	-2.445	0.000	0.087	0.053	0.143
Ethnicity					
White	-4.111	0.000	0.016	0.005	0.052
Black	-2.173	0.002	0.114	0.029	0.443
Native	2.407	0.001	11.104	2.797	44.075
Polynesian	-4.349	0.000	0.013	0.004	0.045
Asian	-3.048	0.000	0.047	0.014	0.158
Other	-2.330	0.000	0.097	0.030	0.313
Multiple	-6.195	0.000	0.002	0.001	0.007
Hispanic	-0.984	0.000	0.374	0.296	0.473
Household Size	0.236	0.000	1.266	1.241	1.291
Commute Mode					
Single-Occupant Vehicle	-6.991	0.000	0.001	0.000	0.005
Carpool	-7.532	0.000	0.001	0.000	0.003
Transit	-.255	0.353	0.775	0.453	1.326
Walked	2.326	0.000	10.238	5.861	17.882
Bike	6.758	0.000	860.609	321.505	2303.688
Constant	10.531	0.000	37473.482		

As the percentage of men decreases in an area, there is a significant decrease in the probability of angle crashes. Education was also significantly correlated to angle crash probability. Areas with higher percentages of people with less than high school, an associate's degree or a graduate degree are significantly less likely to experience angle crashes while areas with a higher percentage of the population with a high school education have a higher probability of having angle crashes nearby. Areas with high unemployment had significantly lower probability of experiencing angle crashes in the area.

Ethnicity was also correlated to the probability of angle crashes in the area. An increase in the percentage of Native American's was significantly correlated to a higher probability of angle crashes. Census tracts with larger household sizes had a statistically higher probability of angle crashes occurring in the area. For each additional person increase in the average household size (for example increasing from 3 persons per household to 4 persons per household), the likelihood of an angle crash occurring in the census tract increases by 25%.

Commute mode was significant at predicting the likelihood of angle crashes. Locations with higher levels of auto commuting (drive alone or carpool) had a significantly lower probability of angle crashes while areas with higher rates of walking and biking have a higher probability of angle crashes. This could be caused as motorists become distracted by non-motorists in the roadway resulting in a crash with another vehicle.

4.5.2 Single Vehicle Crash

A single vehicle crash is a type of road traffic collision in which only the one vehicle is involved. Included in this category are run-off-road collisions, collisions with fallen rocks or debris in the road, rollover crashes within the roadway, and collisions with animals.

A binary logit regression was run to evaluate the relationship between demographics and single vehicle crashes (See Table 11). The model was statistically significant (X^2 = 449.33) and correctly classified over 82.2% of cases.

Table 11. Demographics and Single Vehicle Crashes

Demographics	B	Sig.	Exp(B)	95% Confidence Interval	
				Lower	Upper
Percent Male	2.991	0.000	19.910	13.894	228.553
Age					
Under14	0.006	0.438	1.006	0.991	1.022
Teens	0.074	0.000	1.077	1.060	1.094
Young Adults	-0.083	0.000	0.920	0.901	0.941
Middle Adults	-0.102	0.000	0.903	0.890	0.916
Late Adults	0.019	0.052	1.019	1.000	1.039
Percent Seniors	-0.071	0.000	0.931	0.912	0.951
Educational Attainment					
Less than High School	-3.275	0.000	0.038	0.022	0.064
High School Graduate	-3.206	0.000	0.041	0.027	0.061
Some College	-3.737	0.000	0.024	0.016	0.036
Associates Degree	-4.289	0.000	0.014	0.006	0.030
Bachelor’s Degree	-2.732	0.000	0.065	0.042	0.101
Graduate or Professional	2.355	0.000	10.539	6.732	16.497
Unemployed	3.688	0.000	39.978	22.592	70.744
Ethnicity					
White	-1.850	0.007	0.157	0.041	0.598
Black	-2.272	0.005	0.103	0.021	0.507
Native	0.246	0.758	1.279	0.267	6.120
Polynesian	-1.421	0.054	0.242	0.057	1.023
Asian	-3.302	0.000	0.037	0.009	0.149
Other	-1.138	0.101	0.321	0.082	1.251
Multiple	0.781	0.304	2.183	0.493	9.661
Hispanic	-0.353	0.016	0.703	0.528	0.935
Household Size	-0.482	0.000	0.618	0.605	0.630
Commute Mode					
Single-Occupant Vehicle	-1.477	0.000	0.228	0.152	0.344
Carpool	0.066	0.784	1.068	0.665	1.716
Transit	-2.088	0.000	0.124	0.067	0.230
Walked	-1.391	0.000	0.249	0.132	0.468
Bike	0.698	0.187	2.010	0.713	5.671
Constant	3.344	0.000	28.337		

The percentage of males and teens in a census tract was significantly correlated to a higher probability of single-vehicle crashes. A higher percentage of young adults, middle adults and seniors was correlated to a reduced probability of these crashes. The percentage of residents with a post graduate degree was significantly correlated to a higher probability of single vehicle crashes, while every other educational level was correlated to a reduced probability. High unemployment was also correlated to an increased probability of single vehicle crashes.

Ethnicity was also significantly correlated to single vehicle crashes. Census tracts with a higher percentage of Whites, Blacks, Asians and Hispanics had a reduced probability of single vehicle crashes. Journey to work data showed that census tracts with higher percentages of single-occupant vehicle commutes, transit ridership and walking have a reduced probability of single vehicle crashes. This evaluation did not specifically identify which fixed objects were involved in each crash, but future research could benefit from a more detailed examination that includes crash details.

4.6 Bicycle and Pedestrian Crashes

Bicycle and pedestrian crashes are of concern as they involve more vulnerable road users and are less likely to cluster by location. Bicycle and pedestrian crashes can be located along corridors or be spread throughout a system. This makes planning safety interventions or designing treatments particularly difficult. Understanding their dispersion relative to demographics and area populations may provide improved understanding of these crashes.

4.6.1 Pedestrian Crashes

First, the 5,520 crashes involving a pedestrian were identified within the crash dataset. Next a binary logit regression was run to evaluate the relationship between demographics and crashes involving a pedestrian (Table 12). The model was statistically significant (X^2 = 124.89) and correctly classified over 98.2% of cases.

The percentage of males in a census tract was significantly correlated to a lower probability of pedestrian crashes. Concomitantly, a larger population of children, young adults, middle adults, and seniors was correlated to an increased probability of pedestrian crashes. Areas with a higher percentage of high school graduates with some college or a bachelor's degree was correlated to a significantly higher probability of pedestrian crashes, as were areas with higher rates of unemployment.

One of the most compelling findings of this analysis was that for each percentage increase in the population of Native Americans, the probability of a pedestrian crash increased by 700%. Tracts with larger households had an increase in the probability of pedestrian crashes as well. Lastly, while all commute modes were significantly correlated to pedestrian crash risk, each percentage increase in bicycle commuting was correlated to a 1,700% increase in the probability of a pedestrian crash occurring in the area.

Table 12. Demographics and Pedestrian Crashes

Demographics	**B**	**Sig.**	**Exp(B)**	**95% Confidence Interval**	
				Lower	**Upper**
Percent Male	-1.497	0.003	0.224	0.084	0.594
Age					
Under14	0.100	0.000	1.105	1.056	1.158
Teens	0.034	0.148	1.034	0.988	1.083
Young Adults	0.107	0.001	1.112	1.044	1.186
Middle Adults	0.051	0.020	1.052	1.008	1.098
Late Adults	-0.007	0.784	0.993	0.941	1.047
Percent Seniors	0.250	0.000	1.283	1.210	1.361
Educational Attainment					
Less than High School	0.545	0.474	1.725	0.388	7.682
High School Graduate	1.557	0.009	4.745	1.466	15.356
Some College	1.955	0.001	7.062	2.279	21.883
Associates Degree	-1.427	0.221	0.240	0.024	2.356
Bachelor’s Degree	1.726	0.005	5.617	1.691	18.661
Graduate or Professional	0.691	0.328	1.996	0.500	7.967
Unemployed	3.141	0.000	23.128	4.632	115.472
Ethnicity					
White	-0.416	0.832	0.660	0.014	31.030
Black	2.033	0.381	7.634	0.081	719.972
Native	6.964	0.002	1057.606	12.964	86278.854
Polynesian	-1.104	0.600	0.332	0.005	20.553
Asian	-0.191	0.925	0.826	0.015	44.576

Other	0.159	0.937	1.172	0.023	60.163
Multiple	-0.265	0.903	0.767	0.011	55.049
Hispanic	0.535	0.197	1.708	0.757	3.852
Household Size	0.224	0.000	1.251	1.167	1.340
Commute Mode					
Single-Occupant Vehicle	3.531	0.000	34.145	9.913	117.608
Carpool	3.269	0.000	26.294	6.324	109.322
Transit	4.168	0.000	64.602	11.435	364.987
Walked	7.065	0.000	1170.489	215.375	6361.205
Bike	17.154	0.000	28.17×10^6	20.44×10^5	38.83×10^7
Constant	-9.306	0.000	0.000		

4.6.2 Bicycle Crashes

Over 4,600 bicycle crashes were identified within the dataset (n=4,627). A binary logit regression model was run to evaluate the relationship between demographics and crashes involving a bicyclist (Table 13). The model was statistically significant (X^2 = 50.580) and correctly classified over 98.5% of cases.

Table 13. Demographics and Bicycle Crashes

Demographics	B	Sig.	Exp(B)	95% Confidence Interval	
				Lower	Upper
Percent Male	-0.829	0.115	0.436	0.156	1.223
Age					
Under14	0.038	0.147	1.039	0.987	1.093
Teens	-0.013	0.607	0.987	0.940	1.037
Young Adults	0.029	0.404	1.029	0.962	1.101
Middle Adults	0.076	0.001	1.079	1.031	1.130
Late Adults	0.000	0.989	1.000	0.944	1.060
Percent Seniors	0.072	0.026	1.075	1.009	1.145
Educational Attainment					
Less than High School	0.678	0.419	1.970	0.381	10.185
High School Graduate	0.349	0.590	1.417	0.399	5.036
Some College	1.138	0.056	3.120	0.970	10.034
Associates Degree	-1.731	0.156	0.177	0.016	1.941
Bachelor’s Degree	2.280	0.000	9.776	2.809	34.024
Graduate or Professional	2.894	0.000	18.074	4.761	68.610

Unemployed	1.102	0.224	3.009	0.509	17.795
Ethnicity					
White	-7.837	0.000	0.000	0.000	0.027
Black	-7.295	0.005	0.001	0.000	0.109
Native	2.226	0.358	9.260	0.081	1061.645
Polynesian	-5.940	0.010	0.003	0.000	0.246
Asian	-7.804	0.000	0.000	0.000	0.032
Other	-5.272	0.017	0.005	0.000	0.394
Multiple	-9.553	0.000	0.000	0.000	0.008
Hispanic	-1.799	0.000	0.165	0.066	0.418
Household Size	0.123	0.002	1.130	1.045	1.223
Commute Mode					
Single-Occupant Vehicle	1.413	0.031	4.107	1.136	14.852
Carpool	1.869	0.014	6.481	1.464	28.693
Transit	4.728	0.000	113.105	18.386	695.777
Walked	2.440	0.007	11.475	1.944	67.723
Bike	17.453	0.000	38.01×10^{6}	27.44×10^{5}	52.64×10^{7}
Constant	0.701	0.763	2.015		

Census tracts with a higher percentage of middle adults (ages 35-49) had a significantly higher probability of bicycle crashes. Likewise, areas with a higher percentage of the population with a bachelor's degree or post graduate degree had a higher probability of bicycle crashes occurring nearby. Changes in the population for all ethnic groups were correlated to changes in the probability of bicycle crashes, however only an increase in the Native American population was correlated to an increased probability of bike crashes. The probability of bicycle crashes is higher in areas with larger households, and for each percentage increase in the population who bikes to work, the probability of bike crashes increases by nearly 1,750%.

4.7 Local versus Through Traffic

One major question raised when evaluating the statistical analysis shown in the sections above, was how to determine if the demographics or the populations were the significant factor, or if there was unmeasurable autocorrelation occurring within the model. For example, if we see significantly more distracted driving in communities with a large Native American population, is it that population itself or some other unmeasured variable that is the culprit. Perhaps Native Americans tend to live in areas with other characteristics that are the correlate. To determine if this is the case, a final analysis was run to determine if the individuals involved in crashes were residents of that local community or if they were simply passing through on their way to work or another destination. To determine this, the zip code and city of the driver of each vehicle involved in the crash was coded and compared to the zip code and city where the crash occurred. While the spatial scale of zip codes and city boundaries are not as fine grain as street address, it was the most appropriate way to evaluate this scenario while maintaining the anonymity of the people involved in the crashes being examined. Both zip code and city were included in order to identify the smallest geographic scale possible. Zip codes and cities are not mutually exclusive. For example, in Salt Lake City there are more than 10 individual zip codes, however Riverton and Bluffdale share a common zip code. The spatial scale of zip codes in also not identical. The zip code map for Salt Lake County is shown in Figure 4 below (www.saltlakecityrealestate.com, 2019).

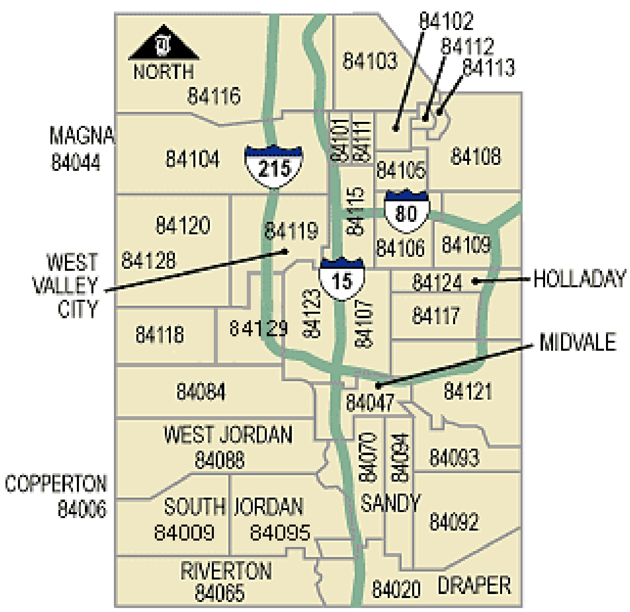

Figure 4. Salt Lake County Zip Codes

First, census tracts were identified based on their significant correlation to a crash characteristic. The crashes occurring within the boundaries of these tracts were then evaluated and the driver’s zip/city was compared against the census tract zip/city where the crash occurred. Each census tract was assigned a value based on the percentage of trips which were local trips (zip codes match) as opposed to through trips (zip codes do not match). This preliminary analysis examined the 408,378 vehicles involved in the crashes included in the dataset. Approximately 76,123 (18.6%) were involved in a crash within their home zip code, while the remaining crashes involved drivers from outside the area.

Figure 5 below shows the percentage of local crashes that occurred within each census tract. It is immediately evident that these local crashes are more common in the southern and western portion of the county. Kearns, Magna, Rose Park and Sandy have a very high percentage of local crashes. West Jordan also has a small pocket of local crashes near Cooper Hills High School, likely due to an overrepresentation of student crashes.

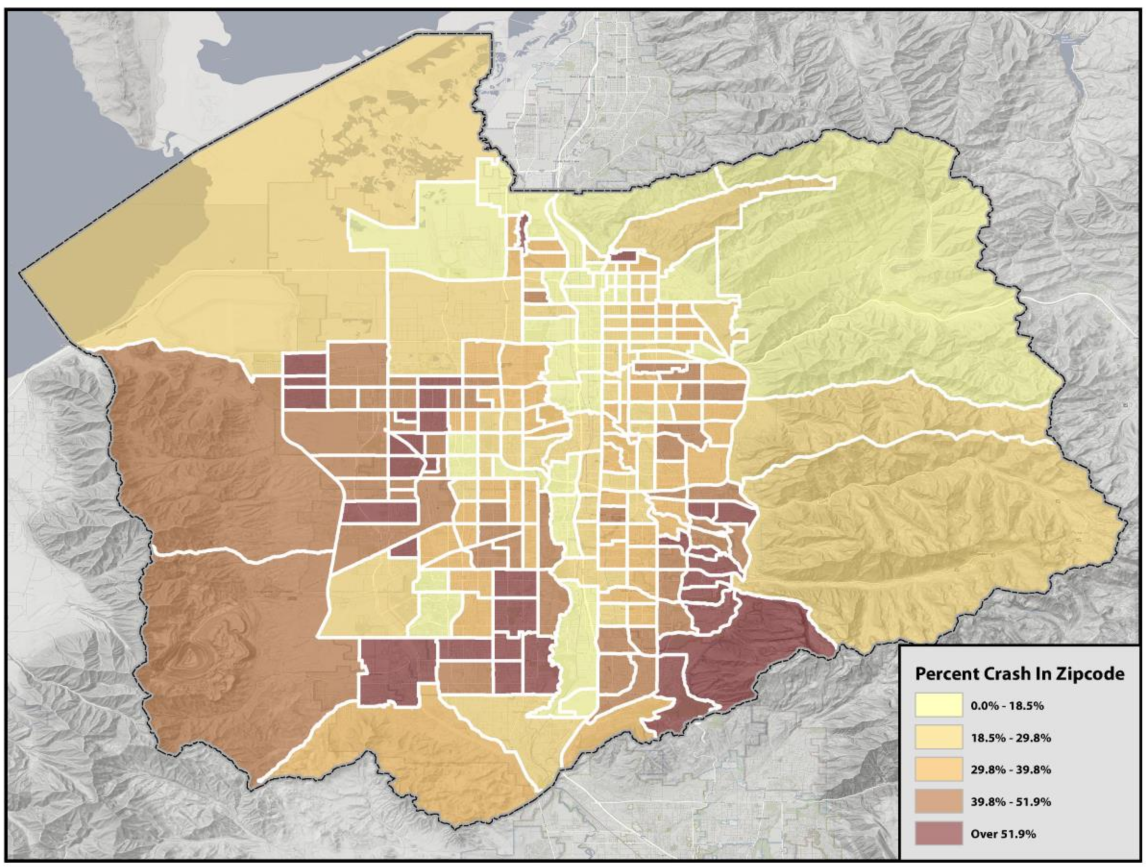

Figure 5. Percentage of Crashes Occurring Within Driver Zip Code

4.8 Summary

The crashes most likely to involve a fatality are those that involve alcohol, crashing into a fixed object, crashes involving a teen or older driver, crashes where the driver or passengers are unrestrained, and those where a traffic control device was disregarded. The most severe injuries occur in crashes involving an angle impact or a single vehicle crash.

Using several binary logistic regression models, demographic factors were correlated to the probability of crash types occurring within a given census tract. Many demographic characteristics were identified with significant correlation. Demographics were also significantly correlated to non-motorized crashes.

Lastly an analysis examined the relationship between driver residential location and crash location. This preliminary analysis found that only 18.6% of crashes occur within the same zip code as the driver's home address, while most crashes involved drivers from outside the area.

5.0 CONCLUSIONS

5.1 Summary

This research examines relationships between demographics and crash risk and rate in Salt Lake County, Utah. Using a combination of sources, population characteristics and other demographic information was identified for each census tract in the county. Additionally, crash data was compiled for all crashes that occurred in the county from 2010-2018. Using a combination of Binary Logistic Regression statistical models, demographic data was correlated to crash characteristics and types to determine: how different crash types cluster across demographic groups and different geographic areas based on population characteristics; and to identify risk factors and criteria for specific populations or areas where populations are concentrated, as well as identifying specific populations which warrant special attention during planning or implementation.

5.2 Findings

The analysis described in the previous chapter show several significant correlations between the probability of specific crash types occurring. In order to ensure that the analysis was robust and thorough, multiple models were run to calibrate the highest goodness-of-fit. In many cases the models revealed high significance across a single category (for example, all age groups were significant). This should not be interpreted as a failure of the models, but rather as a sign of the complexity of the analysis employed. For example, it was determined that household income and percentage of the population with a disability were not significantly correlated to crash risk and type in any of the model iterations. To improve the models, these variables were removed. Additionally, in several cases, age as a percentage was not performing well within the models so an elasticity was created using a scaling effect (low, medium, high) based on the population distribution for each age category (quartiles). These model adjustments were made throughout the analysis in order to optimize the results and outputs. Rather than focus on specific coefficient values in the conclusions section, each crash type will be described with a scale showing the demographic variables that were most significant relative to their impact, for clarity.

5.2.1 Fatal Crash Types

Nearly a third of all fatal crashes involve alcohol (27.4%), which is notable as a very small percentage of total crashes involve alcohol (1.2%). While not typically clustered spatially, several demographic characteristics were correlated to DUI crash risk in an area. As shown in Figure 6 below, areas with a higher percentage of the population with a bachelor's degree and residents biking to work, have a higher probability of DUI crashes occurring in the area. While areas with higher percentages of teens, males, Hispanics, and larger household sizes were negatively correlated. The slide scale shown in Figure 6 provides direct comparison between correlates allowing for easy distinction between which variables were the most strongly correlated. This becomes more important when examining the crash types described in the following paragraphs.

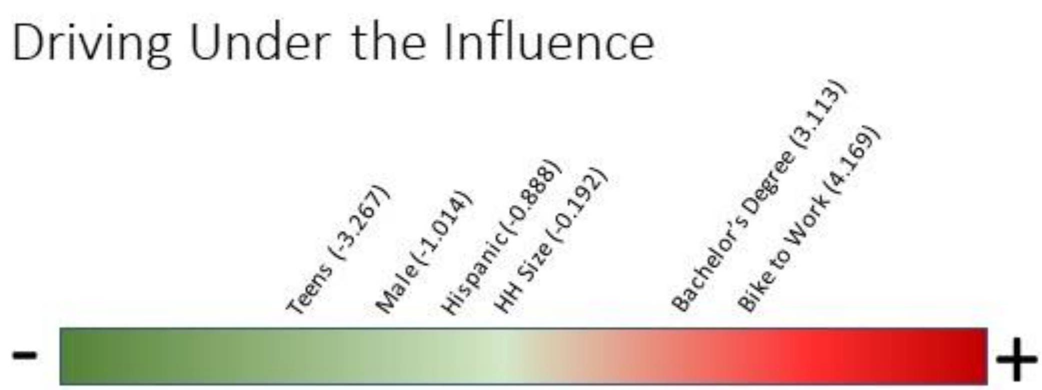

Figure 6. DUI Crashes and Demographics

Because teen drivers are not legally old enough to drink, thus areas with a large percentage of teens would see a lower probability of DUI crashes. Concomitantly, areas where biking to work is most prevalent may have other characteristics that make DUI more prevalent. For example, bike friendly areas tend to have more density, higher access to transit, and be closer to a downtown or urban/suburban center. These would be areas where there may also be a higher prevalence of drinking establishments such as restaurants, nightclubs, and bars. This result opens the potential for autocorrelation, which is described in the following section. Additionally, prior research has shown that individuals with a higher level of education are more likely to be recidivist drunk drivers (Wickens, et al., 2015). It has also been reported that drivers who are involved in a fatal car crash are four times more likely to have a prior DUI than nonimpaired drivers (NHTSA, 2012). Therefore, it follows that areas with a higher percentage of educated individuals would see higher rates of DUI. Figure 7 below shows a map of census tracts based on the percentage of the population with a bachelor's degree.

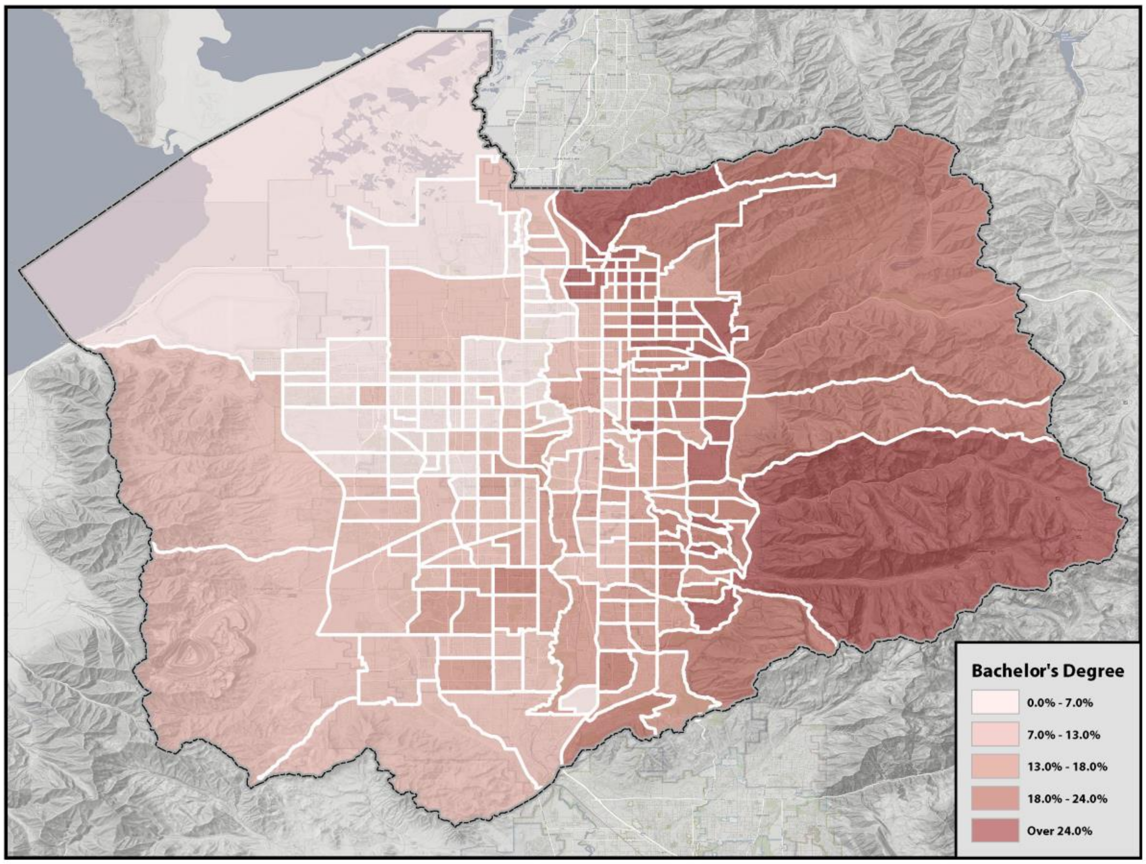

Figure 7. Bachelor's Degree Percentage- by census tract

The areas with the highest percentage of bachelor's degree holders are located on the East side of Salt Lake County. Areas near the Central Business District (CBD in downtown Salt Lake City, the Avenues neighborhood, Millcreek, Holladay, and the area around the Cottonwood Canyons all show over 24% of the population with a Bachelor's degree.

When examining crashes involving older drivers, analysis determined that areas with a higher percentage of males were correlated with a reduced probability of crashes involving older drivers, while areas with a higher percentage of young adults and seniors had a higher probability of these crashes occurring nearby (See Figure 8).

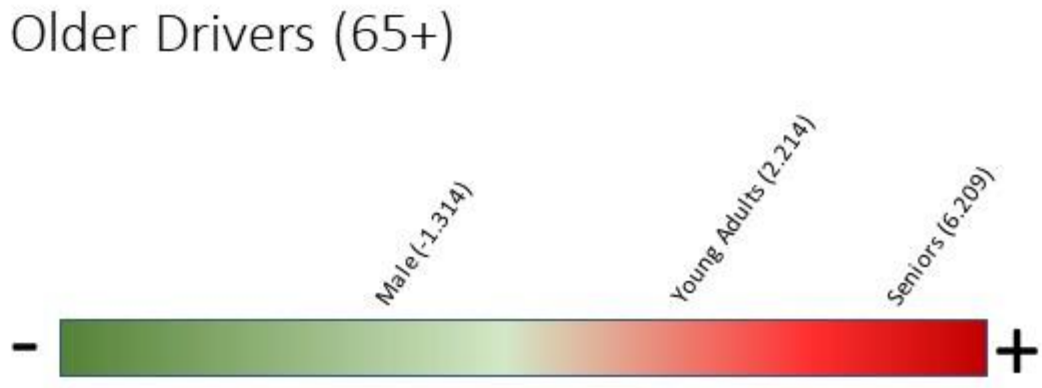

Figure 8. Older Driver Crashes and Age/Sex

The correlation between gender and crashes was seen across nearly all analysis which leads us to believe it has more to do with autocorrelation (discussed in Section 5.3) rather than directly impacting the crash rate. However, each 10% increase in the percentage of Seniors living in a census tract resulted in a 62% increase in the probability that a crash involving an older driver would occur nearby. While we know that most crashes occur outside the river's zip code, the literature has shown that Seniors tend to take shorter trips, staying closer to home than the population at large. As crashes involving older drivers are more likely to involve a fatality, it would be prudent to target areas with large senior populations with increased safety education.

Next, the analysis found significant correlations between older drivers and educational attainment and race. While each category was significantly correlated, it is useful to compare the coefficients and level of impact. Figure 9 below shows that as educational attainment increased, the probability of a crash involving an older driver increased. The probability of a crash involving a senior is nearly 150% higher when the percentage of Bachelor's degrees increases, as compared to an increase in population with less than a high school education.

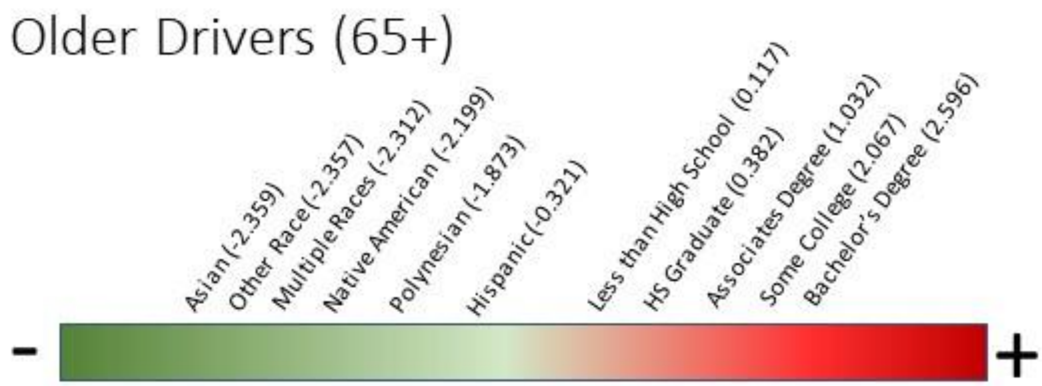

Figure 9. Older Driver Crashes, Education and Race

Areas with a higher percentage of Asia residents were less likely to have crashes involving older drivers which could be due to a cultural tradition where older individuals are less likely to drive themselves, but rather rely on family members. This could also hold true for Native Americans and Polynesians. Journey to work data was also significantly correlated to crashes involving older drivers. The crashes were more likely to occur in areas with a higher rate of alternative commute modes. Figure 10 shows the degree to which commute mode impacts the probability of crashes involving older drivers.

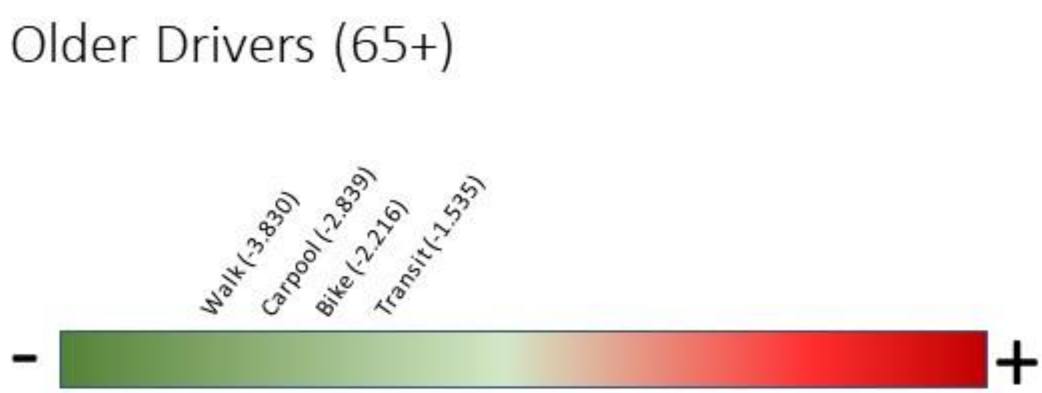

Figure 10. Older Driver Crashes, Commute Mode

Areas with a higher percentage of active transport modes and carpooling were the least likely to exhibit these crashes. While the argument could be made that areas with a higher percentage of seniors would also have lower rates of active mode use, the models controlled for the percentage of seniors, to eliminate the potential for spurious correlations or latent variable impact.

When it came to collision with a fixed object, several factors showed a significant correlation. As shown below, areas with a large percentage of teens had a far lower probability of these crashes than areas with a high percentage of males, or highly educated areas. Again, as an area's educational attainment increased, the probability of fixed object crashes occurring nearby increased.

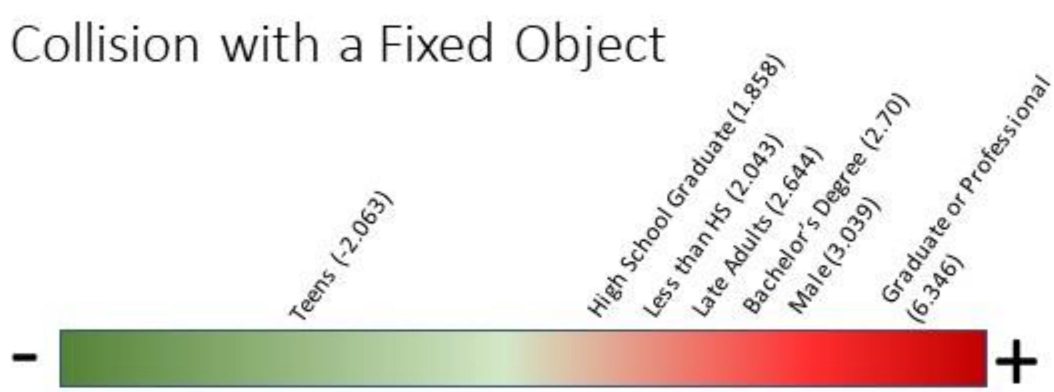

Figure 11. Collision with a Fixed Object- Age, Gender and Education

All ethnicities were correlated to an increase in these crashes occurring except for one. As the percentage of Hispanic in a census tract increased, the probability of a crash with a fixed object occurring decreased (*B*=-2.388).

Journey to work was also strongly correlated to the occurrence of collisions with a fixed object. As the percentage of walking increased in an area the probability of these crashes decreased. While auto-centric commute modes were significantly correlated to a very large increase in the probability of these crashes occurring nearby (Figure 12). It is difficult to speculate as to why these may be correlated, other than to assume that the high rate of auto dominance in commuting may increase the exposure and therefore likelihood that these crashes would be more common, simply due to volumes.

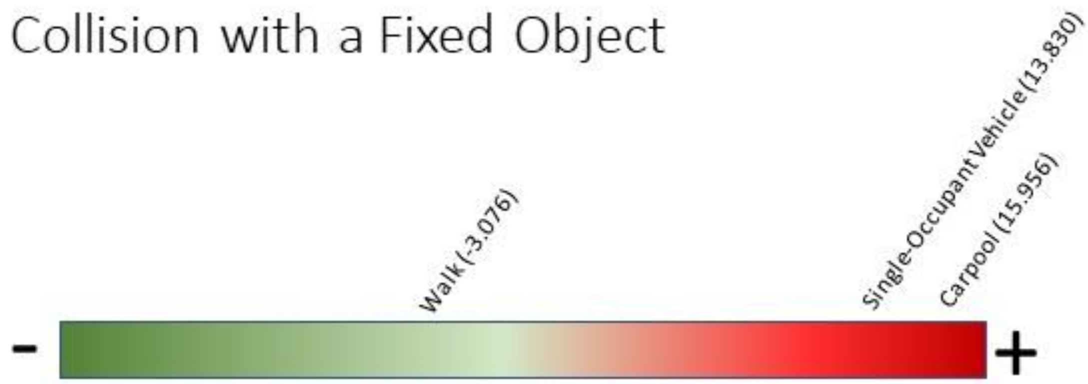

Figure 12. Collision with a Fixed Object and Commute Modes

Next, the analysis of unrestrained crashes showed that age, race, and commute mode were significantly correlated to crash occurrence. As shown in Figure 13, areas with larger households and more Middle Adults (age 35-49) were significantly less likely to experience unrestrained crashes, while areas with more seniors and children under 14 were more likely; but only slightly. Census tracts with a higher percentage of active work trips (Figure 14) and a higher percentage of Native Americans (Figure 15), had the highest probability of having unrestrained crashes occur.

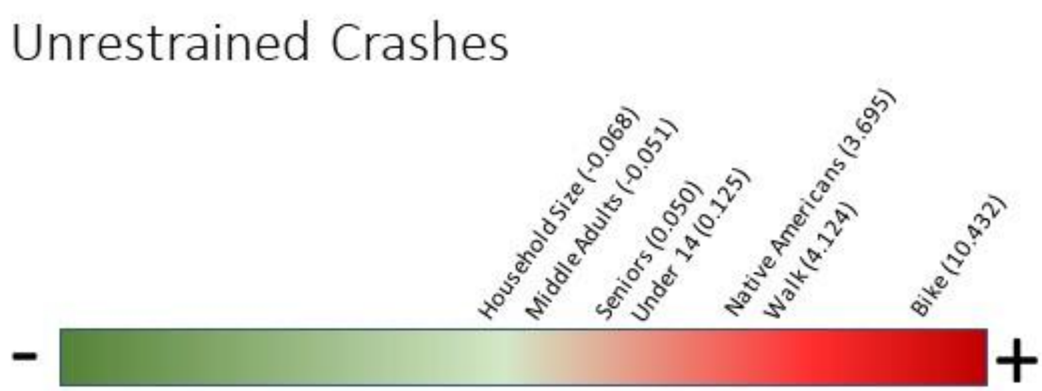

Figure 13. Unrestrained Crashes

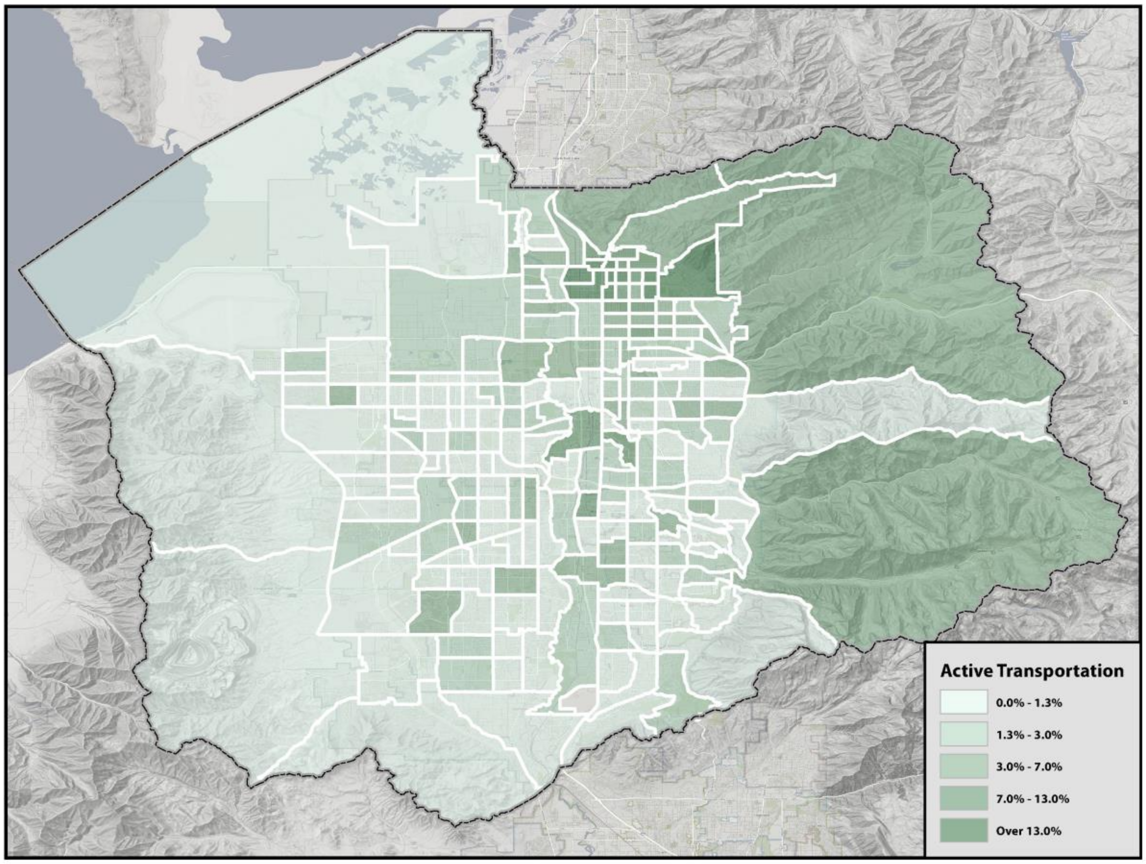

Figure 14. Percentage of Non-Motorized Commute Trips- by census tract

The map shown in Figure 13 shows the percentage of commute trips made by walking or biking. We see the highest concentration of active commute trips near the Central Business District (CBD) in Salt Lake City and near the University of Utah on the East Bench. The east side of the county has a higher portion of active trips than the west side. Figure 15 likewise shows the percentage of the population for Native American residents. These communities are concentrated in several tracts in Salt Lake City, as well as the tract encompassing the Utah State prison population, which would not relationally correlate to crash risk as those individuals are not traveling in the area.

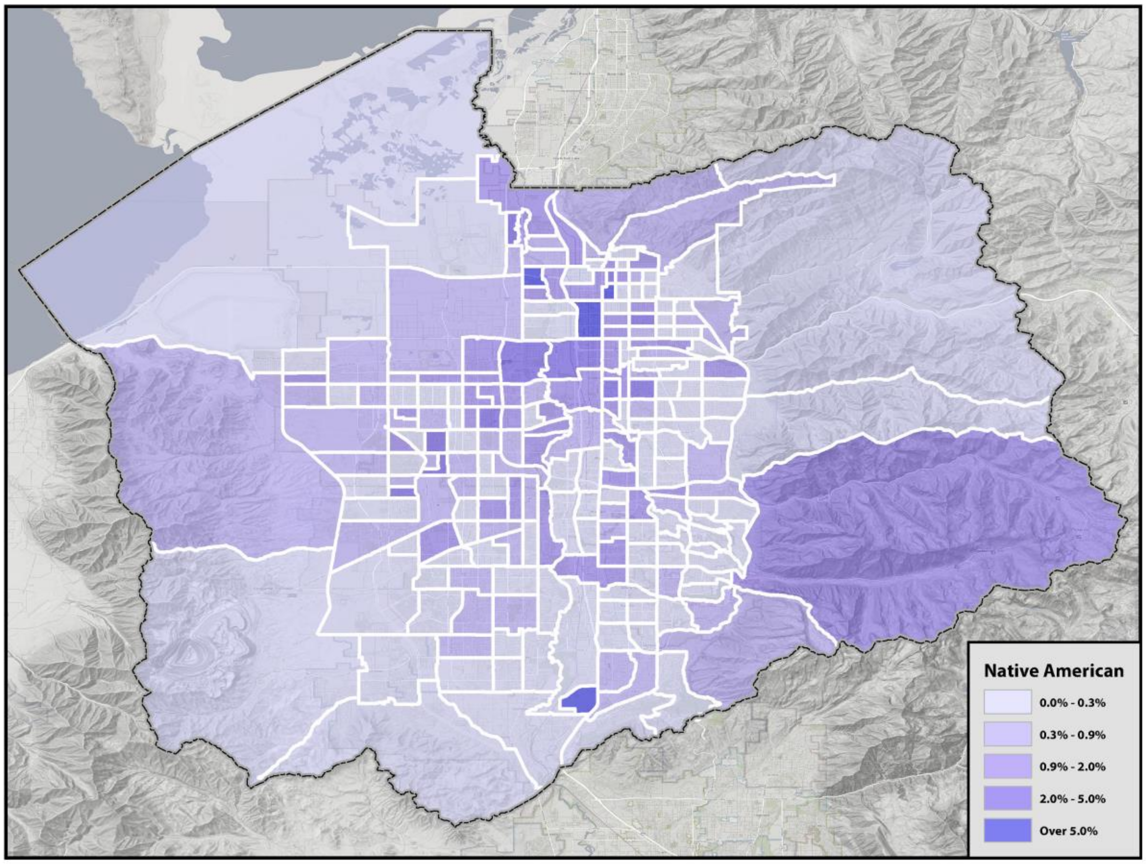

Figure 15. Percentage Native American- by census tract

It is possible that areas with higher active mode share have more low-speed roadways or contain destinations where people would feel more comfortable driving unrestrained. As these areas have lower rates of within zip code it is unlikely that drivers are simply choosing not to buckle up for short trips.

When examining crashes involving a disregard of traffic control devices, gender, age, household size, and commute mode were correlated. Areas with a higher percentage of men have a lower probability of experiencing this type of crash, while areas with higher percentages of children, middle to late adults (age 35-64), and seniors have a slightly increased probability (Figure 16).

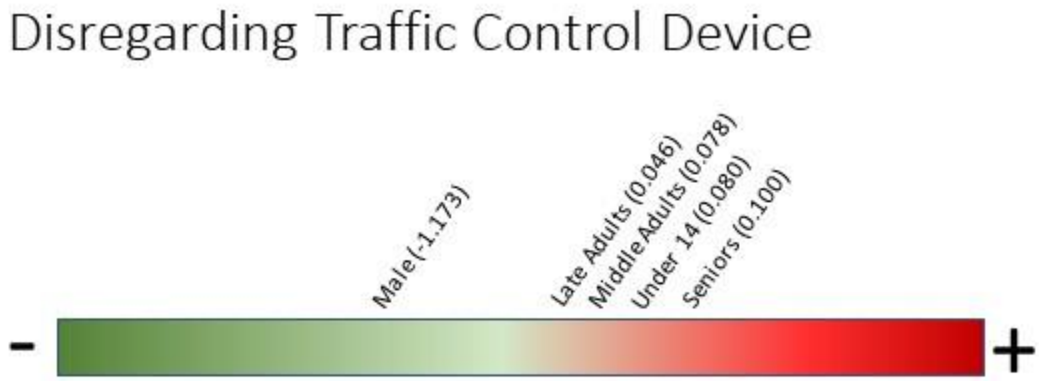

Figure 16. Crashes Involving a Disregard of Traffic Control Device- age and gender

Areas with higher unemployment were significantly less likely to have these types of crashes. Likewise, areas with a highly educated population had a lower probability of crashes involving a disregard of traffic control devices. Census tracts with higher percentages of residents with less than an Associate's degree or a Bachelor's degree had a higher probability of these crashes (Figure 17).

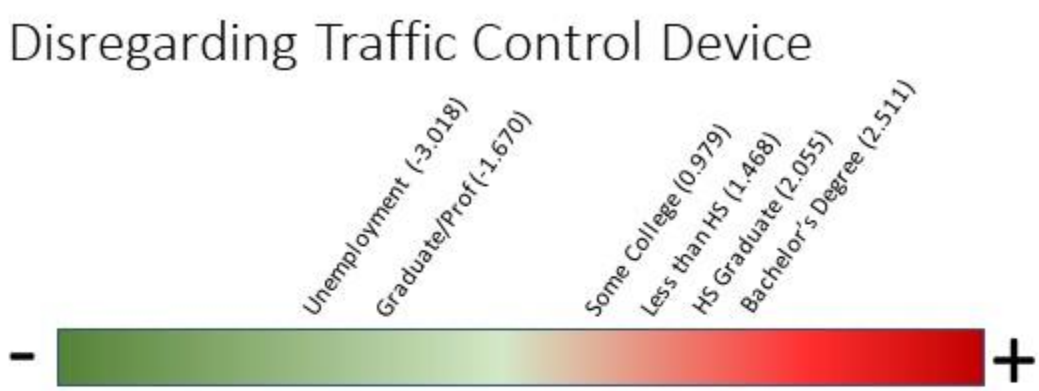

Figure 17. Crashes Involving a Disregard of Traffic Control Device- education/employment

Commute mode was correlated to collisions involving a disregard for traffic control devices. Areas with a higher percentage of transit ridership, walking and cycling had a significantly higher probability of these crashes. In fact, for each 1% increase in bike commute mode share there was a 7.2% increase in the probability of these crashes occurring nearby.

The evaluation of the manner of collision also yielded several significant correlations. Angle crashes were more likely to occur in areas with higher percentages of walking and biking. As non-motorist volumes increase it is possible that this creates a distraction for drivers to observe perpendicular oncoming traffic, as their focus could be on pedestrians or cyclists crossing. Areas with high percentages of driving as a commute mode had a significantly lower probability of experiencing angle crashes. Areas with a high percentage of Native Americans were also significantly more likely to experience angle crashes nearby.

The last evaluation examined correlations between demographics and single vehicle crashes. The data show that areas with a higher percentage of male residents, unemployed residents, or residents with a graduate/professional degree are more likely to experience single vehicle crashes (Figure 18). While areas with a higher percentage of whites, blacks, Asians, and Hispanics had a significantly lower probability of single vehicle crashes nearby. Additionally, census tracts with high rates of transit use and walking had a lower probability of single vehicle crashes.

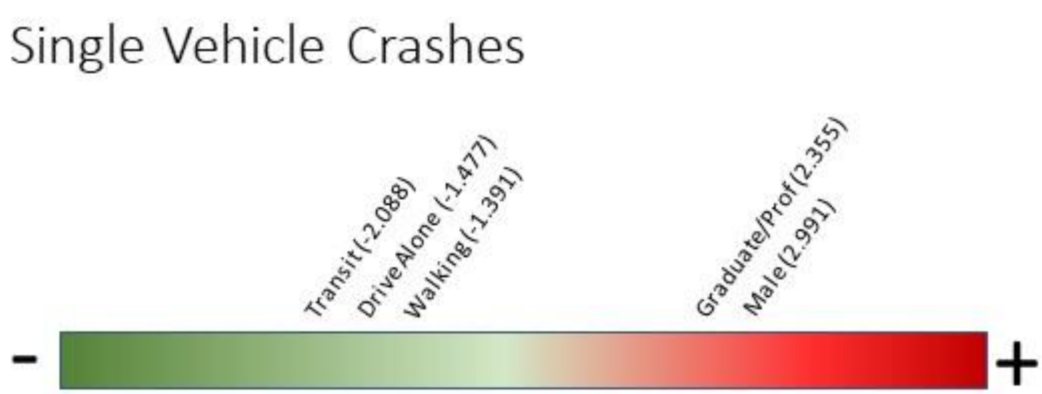

Figure 18. Single Vehicle Crashes

5.2.2 Non-Motorized Crashes

An examination of crashes involving non-motorists (pedestrians and cyclists) determined that several demographic characteristics are correlated. For pedestrian crashes, the most impactful characteristics included employment, ethnicity, and commute mode. Areas with a higher percentage of unemployed residents, Native American Residents, or biking to work was correlated to a significantly higher probability of pedestrian crashes.

Again, this likely encompasses some latent causality with exposure and volumes. As non-motorized travel increases, the exposure to traffic increases which would inherently lead to more crashes involving non-motorists. However, it is important to note that an increase in pedestrian commuters (B=7.065) was not as impactful as the impact of bicycle commuters (B=17.154). Areas with higher unemployment may also inherently experience higher levels of walking and biking due to limited transportation options or an attempt to save money by not driving.

For bicycle crashes, highly educated areas (Bachelor's degree or higher), areas with larger households and areas with higher bicycle commute percentages and transit use were significantly correlated to a higher probability. For each 1% increase in bicycle commute mode share, there was a 17% increase in the probability of a crash involving a bicycle occurring in the area. From the existing literature we know that individuals with higher levels of education are more likely to bike for transportation. This translates directly to the likelihood that bicycle crashes would occur at a higher rate in areas where these populations live.

5.3 Limitations and Challenges

Within the design and scope of this study every effort was made to create a high-quality project that would aptly address the research questions. However, regardless of the care taken every research project will exhibit limitations and challenges. Below is a brief description of the potential limitations present within this study.

First, traffic volume was not accounted for in the models. While this is likely not impactful for vehicular crashes, it likely plays a larger role for crashes involving a bicycle or pedestrian. The analysis found that journey to work percentages within a census tract were significantly correlated to crash risk. In this case, it is probable that journey to work is correlated but is acting as a surrogate for overall volumes. Meaning that areas with more pedestrians and cyclists will typically have a higher probability of experiencing crashes involving those modes. Additionally, data is not collected for the cyclist or pedestrian involved in a crash. It is probable that cyclists and pedestrians use these modes in areas closer to their place of residence as trips tend to be shorter in length. Therefore, the population demographic characteristics surrounding bicycle and pedestrian crashes are more likely to be directly correlated to the individuals involved in these crash types.

Although Salt Lake County has the most diversity in the state, the population is still relatively homogenous. Because of the limited variation in the population, it makes it more difficult for the models to determine strong significance. Additionally, the incredibly large sample size used in the analysis (300,000+ individual crashes) can lead to an over estimation of the models. This means that at some point nearly every variable will show up as significant simply due to the sample size and representation. It is then up to the researcher to explain the nuance and scaling of the model to determine not only which variables are significant, but which are most significant. This was done in the conclusions section by employing a scale of impact rather than focusing solely on the coefficients of each variable shown in the models.

The outputs of the models suggest a small likelihood of serial correlation or similarity between observations as a function of the time lag between them. When error terms from different (usually adjacent) time periods (or cross-section observations) are correlated, the error term is serially correlated. Serial correlation occurs in time-series studies when the errors associated with a given time period carry over into future time periods. This can lead to the conclusion that the parameter estimates are more precise than they really are. Because census data and ACS data are collected over time, and there is built-in estimation and sampling error (as described in Chapter 3) and that error can find its way into the analysis. This problem does not apply to crash data as it occurs over time and does not involve estimation.

Lastly, this study has revealed a high likelihood of latent causation. This means that we are measuring an effect that is caused by a variable not included in the model. Because only 18% of crashes occurred in the zip code where the driver lives (local trips), the local demographics are less likely to be correlated to the crashes themselves, and more likely correlated to environments that exist in those types of neighborhoods. For example, if large household sizes are significantly correlated to a reduced probability of DUI crashes, it may in fact be that areas where larger families live have different built-environment characteristics than the rest of the city (e.g. neighborhoods, further away from larger arterials).

6.0 RECOMMENDATIONS

After a thorough review of the data analysis and findings, the following recommendations have been identified:

- Multiple demographic characteristics were significantly correlated to an increase in the likelihood of specific crash types occurring nearby. As such, UDOT should prioritize employing a thorough evaluation of demographics in areas with higher frequencies of crashes. Traditionally these "hot spot" evaluations look primarily the geometric design of the roadway and may include an examination of surrounding built environment criteria. This research has proven that there are other outside and perhaps even latent factors contributing to the likelihood of crashes occurring and these characteristics should be included in future assessments.

- Trends among specific ethnic communities should be examined further to create appropriate and targeted safety messaging. For example, this research specifically identified that areas with a higher concentration of Native Americans were at risk of having significantly more crashes involving unrestrained drivers and passengers and significantly fewer crashes involving older drivers.

- Journey to work and commute data should be further explored to cross-evaluate which characteristics are increasing risk. This research found that areas with a higher rate of non-motorized commuters had a significantly higher likelihood of experiencing specific types of crashes. Future research should be conducted to determine why this is the case, and if engineering solutions could help to reduce the risk in these areas by better accommodating these active modes outside the flow of motorized traffic.

REFERENCES

Anbarci, N., M. Escaleras, and C.A. Register. (2009). Traffic Fatalities: Do income inequality create an externality? *Canadian Journal of Economics*. January 2009.

Beltramino, J.C. and E. Carrera. (2007). Respect for Traffic Regulations in the City of Santa Fe, Argentina. *Pan American Journal of Public Health*. Special Issue: https://www.scielosp.org/scielo.php?pid=S1020-49892007000700009&script=sci_arttext&tlng=en

Berkley, J. (2017). *Using American Community Survey Estimates and Margins of Error*. Webinar. U.S. Census Bureau Decennial Statistical Studies Division. Available at: https://www.census.gov/content/dam/Census/programs-surveys/acs/guidance/training-presentations/20170419_MOE_Transcript.pdf

Burbidge, S.K. (2016). *Examining the Characteristics of Fatal Pedestrian Crashes*. Utah Department of Transportation. Report No. UT-15.304.

Campos-Outcalt, D. C. Bay, A. Dellapena, and MK. Cota. (2003). Motor Vehicle Crash Fatalities by Race/Ethnicity in Arizona, 1990-96. *Injury Prevention*. (9)3.

De Winter, J.C.F., and D. Dodou. (2010). The Driver Behaviour Questionnaire as a predictor of accidents: A meta-analysis. *Journal of Safety Research*. (41)6. 463-470.

Gardner Policy Institute. (2017). Salt Lake County Small Area Estimates, 2010-2016. Available at: https://gardner.utah.edu/wp-content/uploads/SLCTTractBriefFinal.pdf

Greene, W.H., (2012). Econometric Analysis. 7th Edition. Prentice Hall: Saddle River, NJ.

Keracasu, M., and A. Er. (2011). An Analysis on Distribution of Traffic Faults in Accidents, Based on Driver's Age and Gender: Eskisehir Case. *Procedia- Social and Behavioral Sciences*. (20). 776-785.

Larimer, M.E., G.A. Marlatt, J.S. Baer, L.A. Quigley, A.W. Blume, and E.H. Hawkins. (1998). Harm reduction for alcohol problems: Expanding access to and acceptability of prevention and treatment services. In G.A. Marlatt (Ed.), *Harm reduction: Pragmatic strategies for managing high-risk behaviors* (pp. 69-121). New York, NY, US: The Guilford Press.

Lourens, P.F., J.A.M.M. Vissers, and M. Jessurun. (1999). Annual Mileage, Driving Violations, and Accident Involvement in Relation to Drivers' Sex, Age, and Level of Education. *Accident Analysis and Prevention*. (31)5.593-597.

McDonough, J. and C. Smith. (2010). *Closing the Data Quality Assurance Gap in the Fatality Reporting System*. National Highway Traffic Safety Administration Report No. DOT HS 811 318. Available at: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811318

NHTSA. (2012). *Alcohol-Impaired Driving*. Traffic Safety Facts. U.S. Department of Transportation. Available online: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811870

Oppenheim, I., T. Oron-Gilad, Y. Parmet, and D. Shinar. (2016). Can Traffic Violations be Traced to Gender-Role, Sensation Seeking Demographics and Driving Exposure? *Transportation Research Part F: Psychology and Behavior*. (43). 387-395.

Romano, E. R. Voas, and S. Tippetts. (2006). Stop Sign violations: The role of race and ethnicity on fatal crashes. *Journal of Safety Research*. (37)1. 1-7.

Retting, R.A., S.A. Ferguson, and A.T. McCartt. (2003). A Review of Evidence-Based Traffic Engineering Measures Designed to Reduce Pedestrian-Motor Vehicle Crashes. *American Journal of Public Health*. September 2003.

Roudsari, B. S. Ramisetty-Mikler, and L.A. Rodriguez. Ethnicity, Age, and Trends in Alcohol-Related Driver Fatalities in the United States. *Traffic and Injury Prevention*. (10)5.

Sengoelge, M. L. Laflamme, and Z. El-Khatib. (2018). Ecological Study of Raod Traffic Injuries in the Eastern Mediterranean Region: Country economic level, road user category, and gender perspectives. *BMC Public Health*. (18)236. Available at: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-018-5150-1

Sami, A., A. Najafi, N. Yamini, G. Moafian, M.R. Aghabeigi, K.B. Lankarani, and S.T. Heydari. (2013). Educational Level and Age as Contributing Factors to Road Traffic Accidents. *Chinese Journal of Traumatology*. (16)5. 281-285.

U.S. Census. (2017). *American Community Survey Accuracy of the Data*. Available at: https://www2.census.gov/programs-surveys/acs/tech_docs/accuracy/ACS_Accuracy_of_Data_2017.pdf?#

Wickins, C.M., R, Flam-Zalcman, R.E. Mann, G. Stoduto, C Docherty, and RK. Thomas. (2016). Characteristics and Predictors of Recidivist Drink-Drivers. *Traffic Injury Prevention*. June 2016: 564-572.

World Health Organization. (2018). *Global Status Report on Road Safety 2018*. Available at: https://apps.who.int/iris/bitstream/handle/10665/277370/WHO-NMH-NVI-18.20-eng.pdf?ua=1